



**INFORMATION TECHNOLOGY AND THE EVOLUTION OF THE
LIBRARY**

THESIS

Alexander N. Constantine, Captain, USAF

AFIT/GRD/ENV/09-M01

**DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY**

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government.

AFIT/GRD/ENV/09-M01

INFORMATION TECHNOLOGY AND THE EVOLUTION OF THE LIBRARY

THESIS

Presented to the Faculty

Department of Systems and Engineering Management

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degrees of Master of Science in Research and Development Management and Systems

Engineering

Alexander N. Constantine

Captain, USAF

March 2009

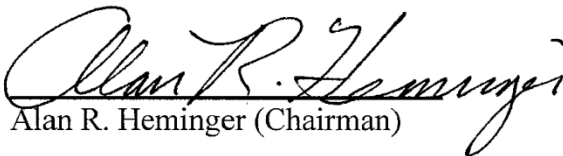
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

INFORMATION TECHNOLOGY AND THE EVOLUTION OF THE LIBRARY

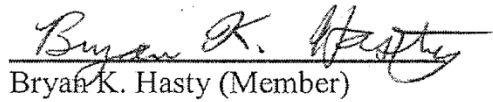
Alexander N. Constantine

Captain, USAF

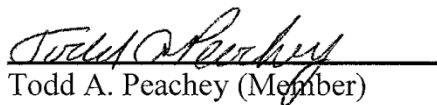
Approved:


Alan R. Heminger (Chairman)

10 Mar 09
Date


Bryan K. Hasty (Member)

10th March 2009
Date


Todd A. Peachey (Member)

10mar09
Date

ABSTRACT

The purpose of this research is to examine the current state of the practice of information technology use among academic and research libraries in order to gain insight into the level of maturity of various emerging technologies such as the Institutional Repository (IR) to be used to provide descriptive guidance and context to other institutions considering them for application. This research uses a case study methodology involving a single case design with multiple units of analysis, and draws evidence from individual interviews of research and education libraries chosen from a range of sizes, foci, and jurisdictions. This evidence is coupled with review of those institutions' public documentation and that of their partners to assess what academic libraries are doing today to meet the needs of the institutions they serve, what they have planned to provide these services, and what unplanned activities they feel are important to continue to make the information they steward available. Respondents cited several key information technology applications including institutional repositories, federated searching, and other integrated library services as tools being migrated towards to help them fulfill their missions. Interview and archival search results demonstrate a marked drive toward increased collaboration among institutions and regional repositories, a desire to provide rich content and seamless interfaces to library customers, and a trend towards hybrid systems to meet institutional archive and faculty communication needs.

DEDICATION

To anyone who ever used a typewriter and stencil duplicator
to publish a thesis,
I salute you.

ANC 2009

ACKNOWLEDGEMENTS

I would first like to thank all those who helped me in my data gathering either through sitting through interviews or by looking over the instrument and providing feedback, especially Eleanor Uhlinger, Doris Kight, George Goncalves, Valerie Moulé, Louis Algaze, Margaret Groner, John Davison, Jane Wildermuth, Dr. Ron Bergquist, Carolyn Hank, Jennifer DeMaro, Lynne Grigsby, Nancy Haritatos, and Dr. Judith Conner. Thanks also to Barry Boettcher and Donna Ericson for your patience and for being such a great pleasure to work with.

Thanks to my advisor and committee: Dr. Alan Heminger, Maj Todd Peachey, and Maj Bryan Hasty for all your excellent guidance and wisdom, to LtCol Patrick Kee for his invaluable help in wrapping up this work and preparing it for its defense, and to the reader for plodding your way through the following pages.

Finally, I extend my deepest gratitude to Nancy Harray for teaching me everything I ever learned about writing a proper paper, and to my parents' forbearance in putting up with 18 months of griping whilst I designed and arranged the following work, without which I'd have found this undertaking quite impossible to manage.

TABLE OF CONTENTS

	Page
Abstract	iv
Dedication.....	v
Acknowledgements	vi
Table of Contents.....	vii
List of Tables	xi
1 Introduction.....	1
1.1 What is the problem?.....	2
1.2 Why is the problem important?	4
1.2.1 New Tools, New Methods.....	4
1.2.2 Competition over Resources.....	5
1.2.3 Overload, Selectivity, and Veracity	6
1.3 What is there to be known about KM in Academic Library Settings?	8
1.4 Towards Discovering Promising Solutions	9
2 Literature Review	10
2.1 Knowledge	11
2.1.1 Knowledge Perspectives & Taxonomies	11
2.2 Knowledge Management & Knowledge Management Systems.....	15

	Page
2.2.1 KMS Success Factors.....	15
2.3 The Library	19
2.3.1 The Library as Knowledge Custodian	19
2.3.2 The Library as a Physical Space	26
2.4 Institutional Repositories.....	28
2.4.1 Implementing Institutional Repositories	33
2.4.2 Functions of IRs	34
2.4.3 Institutional Repositories vs. Journals.....	38
2.5 The Role of the Librarian	43
2.5.1 Librarians as Knowledge Brokers.....	43
2.5.2 The Librarian's Role in Institutional Repositories.....	44
3 Methodology.....	53
3.1 The Case Study Method.....	54
3.2 Data Collection.....	54
3.2.1 The Interview.....	55
3.2.2 Ancillary Data	56
3.3 Sample Selection.....	56
3.4 Data Analysis.....	57
3.4.1 Content Analysis.....	58

	Page
3.4.2 Mapping Methodology Elements to Research Questions.....	58
3.5 Validity	59
3.5.1 Construct Validity.....	59
3.5.2 Internal Validity.....	59
3.5.3 External Validity.....	60
3.5.4 Reliability	60
3.6 Summary.....	61
4 Results.....	62
4.1 The Sample.....	62
4.2 The Questionnaire	63
4.3 Results	64
4.3.1 Library Classification	65
4.3.2 Library IT Initiatives.....	68
4.3.3 Library Unplanned Activities.....	78
4.4 Themes and Trends by Category.....	82
4.4.1 Focus.....	82
4.4.2 Jurisdiction.....	83
4.4.3 Institution Size.....	83
4.4.4 All Categories.....	84

	Page
5 Conclusions and Limitations.....	87
5.1 Conclusions.....	87
5.1.1 What are Libraries Doing Today?.....	87
5.1.2 What have Libraries Planned?.....	88
5.1.3 What Unplanned Activities do Libraries Find Important?	90
5.2 Limitations and Considerations for Future Research	92
Appendix I: Library Interview Questionnaire	96
Appendix II: Summary Response Tables	103
References	113
Vita	120

LIST OF TABLES

	Page
Table 3-1: Mapping research methods to library activities.....	58
Table 4-1: Institution Categorization.....	66
Table A-1: Summary Responses to Classification Questions I	103
Table A-2: Summary Responses to Classification Questions II	103
Table A-3: Summary Responses to Classification Questions III.....	104
Table A-4: Responses to Library IT Initiative Questions I.....	105
Table A-5: Responses to Library IT Initiative Questions II.....	106
Table A-6: Responses to Library IT Initiative Questions III.....	107
Table A-7: Responses to Library Planning Questions I	107
Table A-8: Responses to Library Planning Questions II	108
Table A-9: Responses to Library Planning Questions III.....	108
Table A-10:: Responses to Potential Applications I.....	108
Table A-11: Responses to Library Planning Questions II	109
Table A- 12: Responses to Library Planning Questions III	109
Table A-13: Responses to IR-Related Questions I	110
Table A-14: Responses to IR-Related Questions II	111
Table A-15: Responses to IR-Related Questions III	112

INFORMATION TECHNOLOGY AND THE EVOLUTION OF THE LIBRARY



1 INTRODUCTION

Universities have long been the seat of high learning and expansion of societal knowledge. This knowledge is worthless without the ability to record and retransmit it to those who would build upon it and continue to advance the diverse fields that comprise the human experience. The role of caretaker of such knowledge has been vested in the *Library*. As global society continues to embrace the power of electronic communications in organizing and distributing data, information, and knowledge, the adequacy of physical stores of information would appear to become limited in the context of knowledge transfer, and likewise the efficacy of institutions relying on such physical means to store and transfer knowledge is diminished. While this does not obviate the library or archive as a powerful source of recorded knowledge, the utility of such an entity can be greatly enhanced through embracing the digital and developing a robust knowledge management system whereby the corporate/institutional knowledge of the organization is made available almost instantaneously across the globe. This in turn allows for more rapid evolution of knowledge through participation by broader segments of the population than could ever physically enter any library. In that vein, one might reasonably suggest that *libraries and the institutions they serve can maintain their place as the repositories of knowledge through the development of systems that share their intellectual capital with the widest appropriate audience*. Socrates said that the unexamined life is not worth living (Plato, w. 399 BC; t. aD 1997, p. 54); by extending this idea to living processes like knowledge one might say

that knowledge is irrelevant if it is not shared or experienced in some way, and the institutions that possess it become equally irrelevant. Thus powerful and wide-reaching knowledge distribution systems are an imperative for these institutions to ward off irrelevancy in the changing landscape of information exchange.

1.1 What is the problem?

Anecdote demonstrates that in recent years many who desire to locate information have begun their search with search engines such as *Google*, and more often than not an article in *Wikipedia* or a similar online knowledge source will be in the first several hits. Indeed, so prolific is the use of Google that it has been entered into many English dictionaries as a verb (“I ‘Googled’ it”) (Webster's New Millennium™ Dictionary of English, Preview Edition (v 0.9.7), 2008), and the veracity of the information gleaned from these online knowledge sources is so well accepted, that in casual use many go no further than the first apparently reasonable result that provides them the information or knowledge they desire. The speed and ease afforded by these methods allows knowledge seekers to rapidly generate results to their queries without ever leaving their homes, and reduces their need for locating and perusing dusty tomes for knowledge that is readily available online.

Exacerbating this issue is the fact that as the core of human knowledge and experience continues to grow, so too do the requirements for storage and organization of that knowledge. If we assume that, in general, more knowledge is being generated than destroyed, storing this knowledge with physical (i.e. print) methods is becoming increasingly inadequate – especially for ephemeral, rapidly developing knowledge areas – in-

asmuch as the physical space required to store books and other documents must grow with the number produced, and further as the number of volumes increases, the speed and efficiency with which they can be accessed decreases, no matter how well developed the organization system. Additionally, the costs of producing print sources are not negligible, nor are the costs of the raw materials which must be stockpiled to create a volume, or even the floor space these stockpiles occupy. Electronic media on the other hand provide several advantages in cost, in that once the initial capital has been invested to obtain such a system, the cost to write and rewrite data is essentially equal to the power required to operate it, and the physical size of the system remains constant no matter how much data is stored within (up to its capacity). Additionally, with every advance in storage technology, the physical space required to store the same amount of data actually decreases, as anyone who has ever worked in a large computing center can easily appreciate, (a 1cm×1cm 1G SD-micro card in a digital camera stores more data than 500,000 3.5" floppies could in the 1990s). All of these factors contribute to the decisions made daily in academia, government, and industry to reduce the amount of physical documents produced in favor of electronic sources.

Given the trends discussed briefly above, how does the library fit into the equation? As very physical icons of physical data storage, are libraries becoming anachronistic in the face of the sweeping changes in information technology, or do they provide something more valuable than a simple storage locker for recorded data? Ultimately, libraries can no longer be what they once were and remain relevant in today's climate. To gain insight into the shape libraries will take in the future, we will examine how they are responding to the new questions engendered from the explosion of new technology. This

research will explore the following: (1) *What are libraries doing today to address and incorporate new technologies in serving their customers*, (2) *what have they planned already*, and (3) *beyond what they are doing or what they are planning to do, what do they see on the horizon that may be important to better serve their customers and make the information they steward available?*

1.2 Why is the problem important?

Before launching an investigation into the above questions it is important to discuss why the problem warrants a solution, to determine whether or not an effort should be made, and ultimately to design a framework within which to answer the research questions. Libraries have a large clientele they wish to serve, and advances in technology necessitate changes in the manner in which libraries render that service. In order for any library to explore a new approach to the way it serves its customers, it is important for it to know what approaches are being undertaken by other libraries, what technologies and approaches comprise the state of the art, and what untapped potential exists that it could use to the benefit of the people the library serves.

1.2.1 New Tools, New Methods

A very cursory answer to why the problem is important is simply that with modern advances in information technology (IT), academic institutions have at their disposal a very large new set of tools, and with those tools come new ways of doing business. When new technology “invades” a well-defined, structured industry, the most successful organizations first attempt to gain an understanding of the new technology and the way in which it affects their operations, and then determine the most effective way to incorporate or combat that technology by adjusting their operations to prevent competition

from overtaking them (Utterback, 1994, pp. 162-164). Additionally, as new technology is introduced, successful organizations ponder ways of using it to do things they are not already doing (Hammer & Champy, 2003, p. 89), expanding services into areas that might not have been thitherto in their purview. Ignoring an innovation that could serve to improve the way in which an organization does business is analogous to burying one's head in the sand, and will do little to improve the organization's ability to continue to succeed in the light of a changing competitive landscape.

1.2.2 Competition over Resources

The modern climate of competition in a resource-constrained environment is one not only of competition between institutions, but also within institutions. As large organizations struggle to prioritize budgetary allocations between diverse portfolios of programs, they are forced to evaluate and prioritize their various operations to determine which should receive funding and which should not. How then can an operation such as an academic library reduce the risk of its being affected by institutional cost-cutting? If a library is able to demonstrate that it is heavily used and a vital component of the larger institution's operations, it can more easily ward off the headsman's axe. How does the library increase its usage? The easiest way to do this is by enhancing its relevance and utility to its users. The successful incorporation of new technology, services, and new ways of providing services to users will help maintain the position of the academic library as an invaluable resource and help ensure that it continues to receive the funding it needs to pursue its mission.

1.2.3 Overload, Selectivity, and Veracity

What, then, of the services academic libraries provide? Why and how do they remain essential in an age when information is so readily available? There is a truly staggering volume of data, information, and knowledge available from multitudinous sources on any topic imaginable (Henczel, 2000, p. 210). Finding answers to popular questions is often easy enough, but sifting through what often amounts to a disorganized pile of information can become a herculean task when one is concerned with discovering a more abstruse or specialized morsel of learning. How, then, can knowledge seekers procure tailored data, information, or knowledge to suit their needs? Traditionally, this has been done through a library. The library is more than a simple repository of dated knowledge; librarians are *knowledge brokers*, or individuals who are able to receive requests from knowledge-seekers, translate those requests, and connect seekers to the knowledge sources they desire (Davenport & Prusak, 1998, p. 29). Libraries specialize in finding knowledge, and can analyze the knowledge-seeker's unique request and generate an individualized response in a manner that no algorithm could possibly hope to match. This also means that a library can make researchers aware of sources they mightn't have even been aware of or would not have otherwise found if left to their own devices, thus the library as a data broker remains a very powerful asset in the ongoing quest for knowledge.

A second valuable service provided by an academic library is that of veracity. In order for a researcher to produce a quality body of work, he or she must use appropriate, vetted sources. The open nature of the internet and the ability of individuals to post whatsoever they please bring with them a conundrum inasmuch as it is very difficult for

a researcher to assess the validity of certain sources even when the text passes a cursory sanity check. Even well-respected open-sources such as Wikipedia are editable by anyone, and though they are often well-reviewed by their responsible contributors, new or recently-edited entries, or those entries on obscure subject matter, may fall between accuracy reviews and could easily mislead a consumer into trusting factually-deficient entries. Libraries in general and academic libraries in particular have been respected knowledge authorities since their establishment in ancient days, and as such the knowledge they provide carries with it an air of legitimacy and respectability not present in the newer, less well-established knowledge and information sources. A researcher requesting a particular book or journal is confident that the work he or she receives from a library will indeed be from the source it purports to be, thus libraries provide a powerful asset in their ability to ensure source and content validity. Additionally, the human to human nature of the librarian-customer interaction allows for a two-way exchange that is not possible through purely electronic media. Librarians gain some insight into what journals are well respected in a particular field through years of interacting with researchers, and thus might be able to suggest alternate sources. Researchers also gain insight through their own experiences which they can use to recommend that the library acquire journals and books not currently held on its “shelves” (be they physical shelves or electronic media), and this two-way exchange thus benefits not only the individual researcher, but all those who follow and continue to interact with the library. Thus the library provides peace of mind to the researcher that the source of gathered research is valid (though not necessarily that the information contained within is accurate), and is a representative member of the body of work to which it is attributed.

1.3 What is there to be known about KM in Academic Library Settings?

The previous section established the role of libraries as the de facto knowledge managers. The preoccupation of this research, then, is to determine methods of expanding this role from print to cyber media, and will involve analysis of approaches taken by other academic and corporate libraries.

The opening pages of the next chapter will comprise a review of extant knowledge management literature, and will identify areas of the field that are the most relevant for the academic library setting. The purpose of this section will be to establish and define relevant KM terms and principles and to orient the discussion towards the production of a viable knowledge management architecture for the AFIT Academic Library.

Following the general discussion of knowledge management principles, the first step will be to identify several universities and knowledge centers that have already undertaken KM initiatives, and the manner in which they did so. Next, it is necessary to review which among these efforts have been deemed successful, and by whom. The perspectives of consumers and producers vary, and thus the appropriate context and perspectives must be incorporated into any such evaluation. Along with considering successes, the research will also attempt to identify initiatives that have failed, with the same perspectives and context. We must then determine what factors influenced the success or failures of these systems, and which of these factors were foreseeable and correctable. Following this, we must examine which of these systems (or class of systems) draws the most likely parallel to the subject AFIT/AFRL library, and which factors are of the greatest concern when developing/implementing a solution. Finally we will discuss

which tools are good candidates for incorporation by the AFIT library in enhancing the experience of its customers.

1.4 Towards Discovering Promising Solutions

The goal of this research is to investigate and identify information technology and knowledge management trends in academic libraries, outline what makes such systems successful or unsuccessful, and draw conclusions from this analysis to use as a tool to aid in the development of a framework or architecture for the AFIT Library's own information-technology efforts. This amounts to modeling the various capabilities currently in use and on the horizon to determine what they have to offer an institution such as the AFIT library, identifying likely candidates, and understanding how other academic libraries have implemented and made these services available to their customers. The following section will review the knowledge management literature to provide the first steps towards answering the questions posed in the first section of this introduction, followed by the development of an instrument with which to assess the way in which selected academic libraries have begun to address new technology. The results of this investigation can then be used to help guide the AFIT/AFRL Library's own efforts in enhancing the services it provides to its customers around the country.



2 LITERATURE REVIEW

The issues facing the library, the management of its collections, and the services it provides its clients are all issues which have been embraced and discussed in the knowledge management literature. From its first murmurings in the 1990's the field of knowledge management has grown as various organizations began to recognize the limitations inherent to traditional modes of knowledge storage, retrieval, and sharing brought on by the huge increase in the generation of new knowledge and the proliferation of information system technology. As firms began to consciously recognize the value of knowledge as a corporate asset (Davenport & Prusak, 1998, p. 12), Knowledge Management (KM) developed as a dedicated field of study developed to help organizations understand and cope with the need to retain and utilize this asset to maintain a competitive edge. The best methods to handle this asset differ between various types of organizations, but the overall thrust of KM is to ensure that the correct knowledge is identified, acquired, retained, and made available to those who need it. The proper handling of institutional knowledge is central to the success of all knowledge-based organizations, even more so to those whose purpose for existing is to retain and provide knowledge to a broad audience (such as libraries), and to that end this review will attempt to glean from the literature the most prevalent methods used in the library context, both technological and non-technological.

Before pressing onward with direct analysis of modern institutional repository efforts, their levels of efficacy, and the factors that spell their successes or failures, we re-

view the literature pertinent to knowledge management and modern institutional repositories. As with any evolving business practice, there is the potential for misapplication or faddish abuse of KM and the perversion of its principles into jargon (Davenport & Prusak, 1998, p. xix) or the blind, bandwagon style initiation of knowledge management systems including institutional repositories (Ware, 2004, p. 121). The key is to understand the underlying concepts and processes KM describes to better define an appropriate framework within which to solve an organization's knowledge-handling issues. What follows is a general presentation of knowledge management terms and principles, exploration of literature specific to institutional repositories in the academic setting, and a review of current and historical trends in libraries.

2.1 Knowledge

Knowledge management literature is rife with definitions and taxonomies describing knowledge and how it compares to data, information, wisdom, and other, less tangible ideas. The next section will discuss some of these definitions in brief as a preamble to investigating what knowledge libraries steward and some of the techniques available for sharing and maintaining that knowledge.

2.1.1 Knowledge Perspectives & Taxonomies

2.1.1.1 *Data, Information, & Knowledge*

As stated above, one can define Knowledge in a number of ways depending on one's perspective. According to Davenport and Prusak, "data is a set of discrete, objective facts about events" (Davenport & Prusak, 1998, p. 2), information is data given purpose and meaning, and knowledge is information mixed with experience, insight, and

context within the human mind. This definition takes the more formal perspective of knowledge as a sum of components, data (raw numbers), information (processed data), and knowledge (personalized information) (Alavi & Leidner, 2001, p. 111), and immediately places valuation through hierarchy. Alavi and Leidner also explore knowledge through the perspective of its potential for use and related value to its steward, the capabilities it provides if properly acted upon, or in its effects on individuals, all of which help to identify ways with which to deal with knowledge relative to the perspective of an organization. For the purposes of the university library, the perspective that immediately jumps out is that of *access to information* (Alavi & Leidner, 2001, pp. 109-111), which is at the heart of the academic library's mission of supporting "the educational, research and consulting programs of the Institute" (Air Force Institute of Technology, Public Affairs Office, 2006).

2.1.1.2 Tacit vs. Explicit

Explicit knowledge is that which is codified – the type of knowledge that is structured, accessible, and which can be communicated by symbolic or audio methods. Tacit knowledge is that knowledge which has not been codified. It is in general unconscious or semiconscious in nature and is very personal and often hard to communicate (Nonaka, 1994). Another distinction between these types of knowledge that becomes of great interest to knowledge organizations is that tacit knowledge captures know-how (Brown & Duguid, 1998). Codified or explicit knowledge is very good at conveying *know-what* – routine processes or general information about various topics can be recorded and replayed to impart procedural knowledge – but it is dispositional *know-how* generated

through experience and context that is of true value to an organization (Brown & Duguid, 1998, p. 95).

Vital to managing tacit and explicit knowledge is developing an understanding of the ways in which these forms of knowledge are transferred and the best ways to go about doing so. Tacit knowledge is best transferred through shared experiences whereby the non-verbal and inexplicable nuances of that knowledge can be transferred through example and context. This transfer is largely informal and is accomplished at a more intimate level through discussion and idea exchange than the regurgitation of recorded fact. Tacit knowledge can also be transferred between individuals of similar backgrounds (as in Communities of Practice (CoPs)), but such exchange requires a fundamental common body of knowledge (Alavi & Leidner, 2001, p. 112). This is where the explicit can be brought into play. As in undergraduate institutions, explicit, codified knowledge can be imparted on individuals through tutelage and instruction in theory, but the true know-how is developed as it was in the middle-ages through hands-on through experimentation and practice with the help of senior practitioners in the relevant field (Hammer, Leonard, & Davenport, 2004). By learning a field's basic theoretical knowledge, individuals gain a basic level of familiarity with that field which will enable them to communicate using the community's lingo and view of the world, and as their level of experience in practicing increases, so too will their ability to comprehend methods of applying that knowledge from the senior members of the CoP.

This distinction is between training (imparting of explicit knowledge for rote memorization of some procedural or general knowledge issue), and education (a holistic approach to the investiture of field knowledge and nurturing and development of the in-

dividual as a knowledge worker or master in that field). A successful organizational strategy balances basic training in organizational practices with education through experience, mentorship, and apprenticeship to truly nurture a knowledge-sharing environment that captures both tacit and explicit knowledge (Hammer, Leonard, & Davenport, 2004, p. 17). The organization must allow for the vital informal methods of knowledge sharing between members without reducing structure to the point that the organization is no longer able to function (Brown & Duguid, 1998, p. 105). This approach must be tailored to each organization and functional area to be truly effective since different communities have different norms associated with their knowledge transfer. The preservation of this knowledge for future generations presents the firm with additional difficulties. Even the most rudimentary IT Knowledge Management systems are very capable at capturing the explicit. Tagging and referencing documentation pertaining to just about anything require some degree of effort to make the information useful and effective but in the end amount to developing of a logical organizational structure and efficient search and retrieval system. To capitalize on tacit knowledge, however, a greater degree of interactivity is necessary. As with the organizational structure itself, a balanced KM system is flexible enough to accommodate exchanges within and between CoPs as well as the storage and retrieval of documentation. Thus, for an organization to be truly effective in implementing a KM strategy, it must consider formal and informal methods of transferring tacit and explicit knowledge and develop an environment conducive to both.

Traditionally, libraries focus more on the transfer of explicit knowledge, information, and data in the form of books, journals, and other written sources to knowledge us-

ers. Some technologies, such as video, are capable of codifying tacit knowledge to some extent as well, and may be provided through electronic means to a broad audience. But what happens when the topic of interest to a knowledge seeker is so abstruse or unprecedented that there are insufficient sources available at that library to satisfy the knowledge seeker? In these instances, the academic library can behave as a powerful tool for transferring some forms of tacit knowledge. Researchers may need more than printed work – they may need experts, and the academic library can provide the link between the seeker and the expert. In this way we move beyond content towards providing a bridge for intellectual exchange, and the library evolves from purveyor of documents to purveyor of understanding.

2.2 Knowledge Management & Knowledge Management Systems

2.2.1 KMS Success Factors

Alavi and Leidner define knowledge management systems as “IT-Based systems developed to support and enhance the organizational process of knowledge creation, storage/retrieval, transfer, and application” (Alavi & Leidner, 2001, p. 114). While knowledge management does not require information technology to succeed, information systems can augment the generation, transfer, storage, and accessibility of knowledge and extend the geographic reach of that knowledge and the speed with which it is created and disseminated. Knowledge management technology is available and relatively easy to implement, what is *far* more challenging is overcoming the cultural barriers to contribution to a knowledge management system (Ware, 2004, p. 116), and thus the success of a knowledge management initiative is as much a social effort as it is technical.

This section will focus on a specific form of KMS, the Virtual Community of Practice (VCoPs), which shares many of the key attributes with institutional repositories as will be described in the next section.

Knowledge Transfer is motivated largely through an individual's intrinsic motivation, which may be augmented or hampered by an organization or community's structure and environment (Osterloh & Fey, 2000). Successful Virtual CoPs (VCoPs) require motivated contributors, a supportive organizational climate, and active participation. There must be a supply and demand for knowledge and comfort with the format (Ardichvili, Page, & Wentling, 2003, p. 65).

2.2.1.1 Trust

Trust, or lack thereof, is one of the most important inhibitors to knowledge transfer (Davenport & Prusak, 1998, pp. 96-97), and developing and maintaining trust is essential in encouraging VCoP interaction. Users must feel that contributions will not be misused and that new knowledge is accurate and reliable. This is done through *knowledge-based trust*; that generated through repeated personal interaction, and *institution-based trust*; where a firm's climate makes members confident that others are trustworthy and competent (Ardichvili, Page, & Wentling, 2003, pp. 72-74).

2.2.1.2 Motivation & Participation

In virtual settings, "just-in-time knowledge sharing and transfer" can be accomplished through markets or communities. In discussion forums individuals are likely to provide information to users even without a-priori relationships, inferring a sense of community. The idea of knowledge markets connotes transaction through exchange of

knowledge using some reward mechanisms to motivate behavior (Hahn, Kannan, & Zhang, 2004, pp. 4-5).

Markets operate through supply and demand and are governed by *price*, firms operate through procedure and rules and are governed by *authority*, and communities operate through common beliefs and values and are governed by *trust*. Markets are effective so long as the transaction cost (costs of making exchanges beyond cost to produce the good) is low. When these costs rise firms are better able to handle exchange by creating structure. When assets become too specific or too unique firms cannot adequately predict rule requirements and so communities become more effective (Hahn, Kannan, & Zhang, 2004, pp. 6-7). Further, the more temporally-removed a system or topic of interest becomes from its point of origin, the less likely it is that a firm will maintain the infrastructure to support it (a fact intuitively obvious to those who work with information technology systems), and it often falls to a community to maintain support for that system or knowledge base, especially when the originating firm disappears. This progression from market to community becomes particularly relevant to researchers delving into more abstruse areas of study, where continuity is maintained through a community of practitioners or academicians.

In their 2003 empirical study on motivation and barriers to contribution in VCoPs, Ardichvili et al observed that the major factors encouraging contribution revolved around the perception of knowledge as a public good. Sharing is motivated extrinsically via community interest and obligation, and intrinsically through the desire to establish expert reputations and give back to the community. VCoPs are used as encyclopedias, problem-solving tools, expert systems, and to manage interest groups. The

benefits are to rear novices, enable isolated units to collaborate, locate best practices and lessons learned, find timely answers, and to generate new knowledge (Ardichvili, Page, & Wentling, 2003, pp. 69,70).

In their examination of the barriers to contribution and consumption, they found that in community settings, contribution was limited through fear of contributing unimportant, irrelevant, or inaccurate information, a fear of losing face or of being unworthy to contribute, intimidation at the prospect of being criticized, and organizational security requirements. Barriers to use were a preference for face-to-face interaction and overload (Ardichvili, Page, & Wentling, 2003, pp. 69-70, 71), where there is so much information to sift through the task of internalizing the information becomes daunting.

Another important faced of CoP interaction is the notion that moving knowledge within an organization is harder than sharing it between external related CoPs (Brown & Duguid, 1998, p. 102) – a salient fact when one considers the multidisciplinary nature of most universities. Translation, brokerage, and boundary objects (objects of interest to multiple groups but viewed differently by each) help forge coordination (Brown & Duguid, 1998, pp. 103-104). Translation and brokerage require that both parties trust the intermediary, and it is occupiers of that trusted position that help facilitate knowledge exchange across communities. Furthermore, technology should not constrain the informality necessary to exchange know-how; it should be interactive and allow for participative learning and reciprocity (Brown & Duguid, 1998, p. 106). The next section will examine one institution that has fulfilled the roles of interdisciplinary boundary-spanner and provider of informal knowledge exchange since ancient days – the library.

2.3 The Library

In exploring the relevance of the library we turn to its evolution over time, and the forces which drove that evolution. These forces demonstrate a strong link between technological advances in scholarship and the need to develop tools to facilitate society's retention of the ever-increasing body of knowledge. This section begins with a discussion of the library's historical role as both knowledge custodian and as a physical space, and returns to the modern with a discussion of some knowledge management perspectives on libraries and their role in institutional repositories.

2.3.1 The Library as Knowledge Custodian

Libraries have existed in one form or another from ancient times, often as warehouses for official documents or religious dogma (i.e. the storing of clay tablets for purposes of record keeping to benefit future transactions as in Ugarit, in ancient Sumer) or as the private collections of wealthy or powerful individuals (Wikimedia Foundation, 2008; Ives, Torrey, & Gordon, 1998, p. 269). The form of library that preoccupies this study, namely the universal or research library, developed a millennium after the first recognized libraries, and was distinct in that it made its business not only the storage of the works already belonging to its founder, but rather to gather all the works of the time into its walls, to make those works available to the literate populace, and to encourage the generation of even greater works. This nuance, grown out the Aristotelian mandate of preserving the products of Hellenic thought for the future (Barker, 1956, p. 49), would lead this library to become the most storied library in history, and in so doing establish a powerful precedent for all such libraries that followed. The discussion begins at the turn of the fourth century B.C, in Ptolemaic Egypt.

2.3.1.1 *The Library of Alexandria*

The library of Alexandria was established in the final decade of the fourth century B.C. by Ptolemy I as a center of “learned research,” and would grow under his successors to become the greatest library of antiquity (Barker, 1956, pp. 49,482). This library, along with several other enduring public undertakings took place at the height of the Hellenistic Age; a period during which Greek influence spread across the Mediterranean as far as the Indus valley, and great social experimentation occurred (Langer, 1972, p. 87). The *polis*, or city state, lost significance at this time, allowing for the expansion of Greek culture across the western world as the borders between city-states became indistinct, and with that expansion grew the need to ensure that the transcription of that culture was done in a way that preserved its integrity.

Ptolemy is well recorded as having had many intellectual interests (MacLeod, 2005, p. 62), and as a student of Aristotle and advocate of his ideals (especially those pertaining to the sanctity of Hellenic culture and the vital necessity of collecting and preserving histories (Barker, 1956, p. 49)), he was set on developing a center of learning to rival that of Athens. So desirous was he of gathering the great intellectual minds in Alexandria to pursue scholarly efforts that he established the Library and Museum both to draw in scholars and as a means to ensure the survival of Greek science and literature for the future (Green, 1993, pp. 84-85).

The Library would come to house the immense body of written works produced during the Classical and Hellenistic periods, as well as recorded versions of earlier oral compositions such as Homer’s Iliad and Odyssey. These periods were unique in that the extent of literacy was broad among the populace due to the simplicity of the newly-

acquired 24-character phonetic alphabet (Chambers, Ancient Greece, 1958, p. 11), and works were composed on written media, rather than through strictly oral methods, requiring physical space to store the legacy rather than word-of-mouth transmission (as in the Bronze Age with Homeric poetry). Further enhancing the proliferation of the written word was the development of papyrus as a medium, which thanks to its light weight and compact size (especially when compared to the baked clay tablets or building walls used previously) enhanced the portability of written works to an extent thitherto unimaginable (Ives, Torrey, & Gordon, 1998, p. 270). In modern terms, the phonetic alphabet and papyrus scrolls represent de facto *standards*, the critical enablers of communication between the individual components of a system, and one of the indicators of the development of professionalism in a given field (Buede, 2000, p. 283).

The Library was the first centralized, universal repository of intellectual knowledge (MacLeod, 2005, p. 3) – thitherto scrolls were maintained in private libraries and state libraries housed official records in the manner of archives. The collection of the Library was made available to anyone who “could read, and wished to learn,” which made it a truly public resource (Green, 1993, pp. 89-91). Ptolemy’s desire to encourage the gathering of scholars towards intellectual pursuits, nurtured by Alexandria’s situation as the center of Mediterranean trade resulted in the Library’s growth into an “industry of learning,” and an evolution beyond a simple “repository of scrolls” (MacLeod, 2005, p. 3).

By the time of Ptolemy III, a royal decree mandated that all texts offloaded from ships in the Alexandrian harbor be confiscated and copied, with copies being returned to the owners and the originals taking up residence in the collection (Green, 1993, p. 89). The collection itself spanned a complex of buildings and is believed to have reached al-

most 500 thousand papyri by the end of the third century B.C. (Chambers, Hanawalt, Rabb, Woloch, & Grew, 1999, p. 91).

The Library was dedicated to and emphasized individual research, but the program included teaching as well (Green, 1993, p. 87). The library also had a dedicated staff which developed into a professional scholarly entity, expert at collecting, analyzing, and organizing incoming works and determining their origins. This structure is illustrative of how the advent of another new technology – profession, which began to manifest itself on the social organization of the entire Hellenistic Age and those that followed – would help determine the shape the Library would assume (Chambers, Hanawalt, Rabb, Woloch, & Grew, 1999, p. 91). The ancient head librarians included some of the most preeminent scientists and scholars in the ancient world, such as Eratosthenes (who calculated the circumference of the Earth to within 1% of its value (Langer, 1972, p. 88)) and Callimachus, the first bibliographer (Kinder & Hilgemann, 1974, p. 71). The scientific method, comparative textual criticism, and systematic indexing and cataloging of volumes developed as tools to establish the veracity of works, such as the monumental achievement of the standardization of the Greek text of Homer under Aristophanes of Byzantium and Aristarchus of Samothrace (Chambers, Hanawalt, Rabb, Woloch, & Grew, 1999, p. 91). Eventually, this mission would become the most important of the Chief Librarian's tasks: the establishment of "sound texts, free from spurious matter such as forged interpolations, and purged of scribal errors made during the process of transmission" (Green, 1993, p. 89).

2.3.1.2 Late Antiquity and the Middle-Ages

Throughout the dark ages, monastic libraries collected and preserved works through hand-copying. In Northern Europe, the newly converted monastics rapidly merged traditional Celtic and Germanic art with the reproduction of the books that had become rare and valuable artifacts of Rome (Rosenwein, 2002, p. 58) to develop the new art of illumination. Additionally, the spread of literacy to these cultures allowed for the written composition of some of their legendary epic poems, such as the Anglo-Saxon *Beowulf*, which thitherto were transmitted orally. The Library of Constantinople became responsible for preserving the last works of antiquity with the final destruction of the Library of Alexandria. Libraries engaged in rigorous exchange of works with visiting scholars to ensure copies survived fires at any one complex. Until the fall of Constantinople library efforts focused more on the preservation of knowledge than on its transmission (a trend still observable on the monasteries of Mt. Athos where books are maintained as precious artifacts rather than shared with the greater populace). Among the innovations of this era were the use of miniscule writing to maximize the real-estate on a page and the development of “letter forms that were quicker to write and easier to read” (Rosenwein, 2002, p. 78) by the Carolingians. Indices and tables of contents (which represent the continued evolution of *metadata*, or data about data, from Callimachus’ bibliography) were developed after the 12th century to enhance the accessibility of knowledge within a book to the reader (Ives, Torrey, & Gordon, 1998, p. 270). With the Renaissance and the growth of the Florentine and other libraries, intellectual exchange was again made a priority, and a Catholic Church interested in ensuring a common body of

knowledge among its clergy established Cathedral schools and seminaries to teach, with the libraries and their coveted knowledge at their hearts.

2.3.1.3 The University and the Library

Thus Western universities evolved from these monastic roots to educate first clergy and eventually laity, with illuminated texts at the heart of the transfer of knowledge. Cathedral schools such as the one at Würzburg combined libraries, scriptoria, and schools for clerics in one place as early as the Carolingian period, (Rosenwein, 2002, p. 78) and even published listings of their holdings. Oxford University, the oldest in the English-speaking world – traces its teaching roots as far back as the 11th century a.D. and by 1355 had already been recognized by King Edward III for its contributions to scholarship and services to the state (The University of Oxford, 2008). Cambridge University began as a congregation of scholars at an ancient Roman trading post in the 13th century, and established the world's oldest continuously publishing press in 1584 (The University of Cambridge, 2009). Harvard University was created around the bequest of its namesake, himself a Cambridge graduate, of his entire library and some of his estate in 1638, demonstrating the inextricable link between the university and its library from its foundations (The University of Cambridge, 2009; President and Fellows of Harvard College, 2007). By the sixteenth century, Oxford and Cambridge were so well respected that their opinions on matters spiritual and secular provided foundational and highly-sought-after grounds for legitimacy, so much so that Henry VIII himself bullied them into accepting his divorce and annulment proposals as well as establishing the nullity of marriage after the death of a spouse which would allow him to marry a fourth time. These opinions would provide the basis for rolling acceptance of Henry's purpose by other in-

ternational universities, and when coupled by an exhaustive search by Richard Croke in the great libraries of Italy, garnered him the support of the major universities across France and Italy which in turn gave him the ability to challenge the papacy itself (Gairdner, 1904, p. 105). This event evinces a marked departure from early-middle age concepts of legitimacy. During the rise of the fledgling Roman successor states, Germanic rulers without the benefit of lengthy dynastic successions sought out papal sanction as kings, thus bolstering their claims to the throne while providing the Church necessary support in uncertain times (Rosenwein, 2002, p. 74). Nowhere was this more evident than with Pope Leo III's coronation of Charlemagne as Augustus and emperor of Rome, which provided Charlemagne's usurping Carolingian Dynasty divine legitimacy, and provided the Pope protection from his rivals in Rome in the form of Charlemagne's army (Einhard, w. ca. AD 770-840; t. AD 1998, p. 33). 736 years later, a willful monarch was able to harness the legitimizing power of the great universities and libraries of the age to directly oppose a well-established system of trust and authority, speaking volumes to the power of the university and libraries as bastions of knowledge and scholarship to present research and opinion that could challenge even the vast power of the Pope and the Holy Roman Emperor.

The development of the printing press in the 1450s by Johann Gutenberg (Rosenwein, 2002, p. 201) represented a huge technological breakthrough in Western Europe, in the same way as literacy had in antiquity. The ability of this device to rapidly generate hitherto unimaginable volumes of text released the monks from their arduous task of hand-copying and illuminating manuscripts, but brought with it new concerns about the credibility and legitimacy of those texts. Whereas before the limited sources

of texts allowed for careful consideration of each book by respected authorities (the monks whose task it was to copy them and the abbots and bishops above them), the ability of any literate individual with the skills to typeset a page to rapidly generate a volume allowed for the proliferation of heretical texts and quickly became a burden to Henry VIII and his bishops. In order to combat the spread of false translations of the New Testament and other heretical works, Henry instructed his bishops to gather the greatest scholars to the university to prepare a new translation of the scripture (Gairdner, 1904, p. 106). Here again, the credibility of the University would challenge popular media and help to establish the source validity and primacy of works in a market flooded with new information at the advent of mass communication – a role that translates directly to the present day with the proliferation of the internet and electronic forms of communication.

2.3.2 The Library as a Physical Space

The modern layout of the library takes its roots from the Lyceum of Aristotle (Chambers, Hanawalt, Rabb, Woloch, & Grew, 1999, p. 80), upon which the Library of Alexandria was modeled (Green, 1993, p. 85). This division of rooms and common spaces was designed to allow not only for storage of written works, but for spaces for contemplation, for teaching, for the open exchange of ideas between scholars, and as a peaceful refuge from the bustle of city life beyond the walls of the building. Modern academic libraries have much the same format; they are designed around areas for quiet reading, carrels for focused studying and work, open areas for congregating, and spaces for group study and review. It remains a place to seek knowledge, to consult experts in seeking knowledge, to interact with peers, and to exchange knowledge and information.

The ancient library was not a mere repository; it was a place where scholars convened to discuss science and philosophy, to meditate, and to produce new knowledge. The exchange of ideas between individuals with and seeking knowledge has been the function of the library from antiquity, and remains so now. All of this shows that the Library is intimately involved with not only the storage of knowledge but its generation and provides an historical mandate for the library to pursue these efforts. The Library grew from the need to maintain the social and intellectual products of classical culture as they grew too large to repose in the individual mind. The survival of texts from antiquity was ensured through exchange and duplication of manuscripts, and comparative criticism evolved to establish the validity of sources and content. The Library was a gathering place for scholars and scientists from around the Hellenistic world, setting the historical precedence for its function as a place where experts were connected to produce greater achievements than were possible by solitary research. Unlike its predecessors, the Library of Alexandria's historical "success" as an institution is tied as much to what it was able to do with its material as to how great an amount of content it could hold, and *that* is the more powerful distinction. The roles and skill sets of librarians and the dedicated staff of experts who manage libraries evolved as new technology was introduced. With the advent of the new possibilities granted by modern information technology, libraries are at the cusp of yet another natural step – by embracing the possibilities offered by new technologies the library can continue in its role of stewarding extant knowledge, gathering the learned, and supporting the generation of new knowledge.

2.4 Institutional Repositories

With the advent of the information age, not only has the volume of intellectual creation increased, but so has the speed of communication of results, forcing those eager to exchange the results of their experimentation and collaborate with other scholars to exchange prepublication works on a one-to-one basis, allowing them to continue working while awaiting official publication in a journal. Compounding this, publication may never come, leaving this so called “grey literature” floating in the ether, never to enter the official record (Smith, 2002). Institutional repositories were established in an effort to harness the emerging capabilities offered by new information technology and represent a reform to the system of scholarly communication, returning control over scholarship to academies, and helping to demonstrate the quality of these institutions and their academic pursuits (Crow, 2002, p. 4).

An Institutional Repository (IR) is defined as “a set of services that a university offers to the members of its *community* for the management and dissemination of digital materials *created by the institution* and its community members” (Lynch, 2003), or as “digital collections capturing and preserving the intellectual output of a single or multi-university *community*” (Crow, 2002, p. 4). The italicized portions of this definition call out two important aspects of IRs that will be discussed later in this section: institutional repositories are developed by and for institutions (or even larger entities such as nations) (Prosser, 2003, p. 168), and that their content is produced by the members of those institutions’ communities. They provide a centralized archive (Gibbons, 2004, p. 12) with decentralized self-publication (or self-archival) by the authors (Allard, Mack, & Feltner-Reichert, 2005, p. 325), and are distinct from the individual websites of scholars,

departmental archives, e-print archives, and disciplinary archives in that they reflect the intellectual capacity of a single organization across a wide variety of media (written, video, and so forth) while encompassing diverse disciplines (Bailey, *The Role of Reference Librarians in Institutional Repositories*, 2005, p. 260). The Scholarly Publishing and Academic Resources Coalition (SPARC) further specifies that a digital IR “can be any collection of digital material, *owned* or *controlled*, or disseminated by a college or university, irrespective of purpose or provenance” (Johnson, 2002, p. 2). Here the words *owned* and *controlled* are italicized to demonstrate that intellectual property issues are one of the great confounders to populating IRs, as the material that is owned and controlled may not necessarily be the same material that the university or college creates.

The interdisciplinary nature of IRs falls well in line with the model of organizational knowledge generation insofar as they force disparate CoPs to interact within one forum, helping to produce the “productive tension” between these disciplines that enables more gestalt thinking (Brown & Duguid, 1998, pp. 97-98). The development of the institutional repository as a distinct form of knowledge management system can trace its origins back to MIT’s DSpace project which entered service in 2002 (Ware, 2004, pp. 115, 116), and earlier to Cornell University’s disciplinary ArXiv.org project which provided content derived from several major hard sciences in the 1990s (Genoni, 2004, p. 301; Ware, 2004, p. 116). The first part of the section will review a few IR efforts stood up by some major universities and academic centers, for the purpose of highlighting certain common attributes of institutional repositories. The discussion will then move to the considerations and challenges involved in implementing institutional repositories, their functions, and competing interests.

DSpace is an open-source Institutional Repository package which began development in 2003-2004 under a grant from the Andrew W. Mellon Foundation and Cambridge-MIT Institute (DSpace Foundation, 2008). It was built by The MIT Libraries and Hewlett-Packard (HP). The original DSpace Federation included Cambridge University, Columbia University, Cornell University, Massachusetts Institute of Technology, Ohio State University, and the Universities of Rochester, Toronto, and Washington, and the DSpace community is defined as the group of institutions using this IR software package to manage their collections. It is designed as an open-source system which may be freely downloaded and modified to suit an organization's particular needs (DSpace Foundation, 2007). DSpace was initially created to challenge the "well-worn paths" of the scholarly communication process whereby only certain summarized results of research are published in prestigious journals (often after lengthy submission and review processes), thereby gaining credibility and respect while leaving the community at large without the greater body of pre- or un-published works which often contain more detail or more current work than that which is published (Smith, 2002, p. 543). This process is anachronistic at its core, derived from a time when journals were the de-facto mode of communicating scholarly achievement. The DSpace system was established as an effort to capture those documents and preserve them, along with many other forms of primary digital data (images, datasets, code, etc), and make them available to be built upon by future scholarly pursuits (Smith, 2002, p. 544). Cornell University's experience implementing DSpace reveals certain barriers – especially cultural – to the full realization of an IR which are instructive in determining an approach to populating a repository with the knowledge the institution desires (Davis & Connolly, 2007, p. 1).

The University of California developed the eScholarship system to provide “low-cost, alternative publication services for the UC community, support widespread distribution of the materials that result from research and teaching at UC, and foster new models of scholarly publishing through development and application of advanced technologies” (The Regents of the University of California, 2008). The ePublishing system supports the University’s mission by incorporating numerous applications to archive and present scholarly work in topical areas and to keep consumers apprised of the status of their areas of interest (i.e. new submissions) through alerts (Ware, 2004, p. 117). The California Digital Library (CDL) was itself established as a distinct entity in 1997 to advance the “assembly and creative use” of the various libraries in the UC system and their service to their communities, and defines its mission as “supporting scholarship, building collections and services, and fostering innovation and collaboration (The Regents of the University of California, 2008). CDL is demonstrative of both the “institutional” and “national” uses of institutional repositories, as it powers not only the repository of the University of California (<http://repositories.cdlib.org/escholarship/>), but also the primary-source (<http://www.calisphere.universityofcalifornia.edu/>) and media (<http://www.oac.cdlib.org/>) collections for the state of California (The Regents of The University of California, 2008), literary archival efforts (<http://www.marktwainproject.org/homepage.html>), and even a public access to Melvyl (<http://melvyl.cdlib.org>), the UC’s venerable library catalog system (The Regents of the University of California, 2006) among other programs all geared towards providing digital access to these precious resources and helping to preserve them. Berkeley also offers an IR-construction service through the Berkeley Electronic Press’ Digital Commons (The

Berkeley Electronic Press™, 1999-2008). This is the service that powers ePublishing's web-submission and distribution system (Ware, 2004) and offers a hosted, tailored IR that matches the owning institution's web format. The system is licensed rather than open-source and provides a full-service alternative to DSpace with several features not present in the open-source system such as peer-review modules and automatic conversion of word documents into PDF formats (The Berkeley Electronic Press™, 1999-2008). The contrast between this system and the open DSpace initiative highlights the competition between market- and community-based approaches to knowledge management described above (though both can be *used* by their owning organizations in an open-source manner), a competition which manifests as one between IRs themselves and paid-subscription journals.

DAEDALUS was developed as part of the *Focus on Access to Institutional Resources* (FAIR) program which aimed to implement an *Open Archive Initiative-Protocol for Meta-Harvesting* (OAI-PMH) compliant repository at the University of Glasgow (Ware, 2004, p. 117). OAI-PMH is a standard that allows third-party *harvesters* such as Google scholar to access different forms of metadata. This exposes this metadata to the web, which allows search engines to access data from across repositories thereby linking all compatible systems in a global network (Shearer, 2003, p. 253), and is an important part of providing access to an IR and ensuring interoperability. The DAEDALUS project made a concerted effort to bolster faculty contribution to its system, especially in regard to journal articles following a period of self-reflection over concern at what the developers of the repository felt to be a relatively low level of content (Mackie, 2004). This experience

provides a good complement to the discussion of Cornell's issues with implementing DSpace in identifying ways to combat the cultural phenomena that endanger IR success.

The SHERPA, or Securing a Hybrid Environment for Research Preservation and Access project is another UK-based FAIR program concerned with promoting and advancing the development of open institutional repositories (SHERPA, University of Nottingham, 2006). Its services include RoMEO, which contains summaries of copyright information for various journals to aid those seeking to incorporate journal articles into IRs; OpenDOAR, a list of global open IRs; JULIET, a list of guiding policies from various research sponsors; and a full-text searchable database of UK repositories; among others. The information contained within these services, especially RoMEO, was useful to the DAEDALUS project when identifying candidate content to bolster its lagging population (Mackie, 2004, p. 2), and is one of a number of services which any library can use in developing a plan to create an institutional repository.

2.4.1 Implementing Institutional Repositories

In many respects, the Institutional Repository is a special case of the more broadly-implemented Virtual Community of Practice (VCoP): IRs have contributors (those who load content), consumers (those who use the content), administrators (those authorized to monitor and manage the system), and subscribers (those recognized by the system and able to interact with it in some way). Institutional repositories exist in both subject and interdisciplinary forms, have categorization schema, and have standards both for acceptable content and conduct (Gibbons, 2004, p. 9). Unlike VCoPs, institutional repositories tend to reside within one organization (and indeed represent that organization rather than a community at large) (Prosser, 2003, p. 168), though like VCoPs

their content is made accessible to a broad audience and there are avenues for interaction between that audience and the contributors. These parallels provide a powerful tool for analyzing the efficacy of these systems and the barriers to their implementation.

2.4.2 Functions of IRs

The functions of institutional repositories may be boiled down into two key elements: availability and preservation. IRs enhance and expand access to an organization's intellectual capital and provide a means of centralizing these works, thus increasing the ability of the organization to preserve them (Allard, Mack, & Feltner-Reichert, 2005, p. 327). They provide their host institutions with several core functions, namely material submission, metadata application, access control, discovery support, distribution, and preservation (Gibbons, 2004, p. 7).

2.4.2.1 *Stewardship and Scholarship*

One of the more immediately recognizable benefits of a centralized digital repository is its ability to preserve the works in its purview. Grey Literature is the body of unpublished work produced by scholars which has come to be one of the principal modes of communication between individual scholars who are unwilling to wait for the formal publication process and its shortcomings (Bell, Foster, & Gibbons, 2005; Gibbons, 2004; Ware, 2004; Genoni, 2004). Grey literature in its purest form is often incomplete and premature, yet this prematurity can be an asset to an organization desirous to encourage innovative thought and open exchange among members of the community because the nescience surrounding it leaves room for development and maturation of an idea in ways not originally envisaged by its author (Gray, 2003, p. 22). This literature encompasses a wide range of media, can include presentations and even raw data, which

can offer future generations a great legacy of work upon which to build with new techniques and ideas should it be preserved, yet it is this form of scholarly produce that is least likely to be published and thus least likely to enter the public record once its creator leaves academia (Genoni, 2004, p. 301). Institutional repositories offer a centralized location for all such literature, which aids in the preservation of that literature for the future as well as opening its access to a broad range of potential collaborators without the need of a great deal of individual effort on the part of the original author, whereas personal web pages require the author to have a modicum of skill in HTML, and more local forms of storage require effort on the part of the author to distribute literature to collaborators (Gibbons, 2004, p. 12).

While the preservative benefits of institutional repositories are relatively undisputed, the library community is divided over the issue of *access*, especially against the backdrop of subscription journals (Davis & Connolly, 2007, p. 2). In many ways, institutional repositories were envisaged as a way to break the monopoly of scholastic journals over academic communication, thus enhancing the ability of scholars to exchange ideas and results without the down-selection, time, and expense inherent to the journal submission process (Crow, 2002, p. 4). Position papers released by members of SPARC including Raym Crow's 2002 paper and David Prosser's 2003 speech advocate removing barriers to scholarly communication to increase the dissemination and impact of individual scholars' work (Prosser, 2003, p. 168). Providing low- or no-barrier access to these products increases the awareness of scholarly contributions. D-Space was created with open access in mind, so much so that following the initial development of the software MIT immediately launched into a program to make the D-Space software itself open-

source (Smith, 2002, pp. 547-548). The open access movement aims at the wholesale removal of these barriers, laying bare scholarly literature across the entire globe and making all scholarly journals free to the public, and Prosser goes so far as to use the term open-access in his definition of the functions of IRs (Prosser, 2003, pp. 168-169). *Open-access* and *open-source* describe a common ethos shared by many academicians and software programmers, respectively, that aims to remove copyright protection as an impediment to the free and unrestricted distribution of intellectual products (Baase, 2008, pp. 235-237; Prosser, 2003, p. 169). This is distinct from free access, where there may still be limiting copyright restrictions which limit the acceptable use of intellectual property in some way – in open access, the only appropriate limiting factor is proper name-attribution for purposes of source validity and allowing authors the ability to control the integrity of their work (Bailey, The Role of Reference Librarians in Institutional Repositories, 2005, p. 262).

While open access repositories offer unfettered distribution of an author's work to the world and effectively remove all barriers to scholarly communication (Johnson, 2002, p. 6), reality may limit an institution's ability to completely open all of its work. Both open source and open access are altruistic enterprises that shift the expense of creating intellectual products away from the consumer to their producers and distributors, which must pay to generate and host these products without the expectation of direct remuneration (Baase, 2008, p. 235; Prosser, 2003, p. 169), and this arrangement could prove expensive for producing institutions. Clifford Lynch deliberately separates out the terms "scholarly communication" and "scholarly publishing" and defines repositories in the latter role, as a means of augmenting and enriching the current communication para-

digm (Lynch, 2003). In his discussion on *the role of reference librarians in institutional repositories*, Charles Bailey notes that IRs and open-access are not synonymous, and that while open access is easy to understand, there are many mitigating factors which confound its execution (Bailey, *The Role of Reference Librarians in Institutional Repositories*, 2005, p. 259). These interposing factors include copyrights not held by contributors, the need for privacy or protection of incomplete works to prevent scooping by members of a community of practice unaffiliated with the particular institution or even competing research teams, or moratoria imposed on institutions by partnering organizations, firms, government agencies, etc to prevent release of competition-sensitive material (Johnson, 2002, p. 5).

Removing barriers can be approached by degrees: institutions can allow the contents of their repositories to be searched by external search engines by exposing metadata to the web. This reduces the complexity involved with creating a search engine to search the repository while exposing only metadata to the internet (Johnson, 2002, p. 5). External search tools such as Google Scholar or OAIster can then “harvest” the metadata from various listed journals and repositories and delivers query results to the seeker (Gibbons, 2004, p. 9). The seeker must still have access to the journal or IR to download the content, thus preserving the integrity of the content and the desires of the owning institution while advertising its existence without a user ever having to log onto a specific IR system. In this way, the system can still define and control access, not only internally (specifying roles of contributors, administrators, and so on), but externally (specifying which users can access content and which cannot).

2.4.3 Institutional Repositories vs. Journals

Scholarly journals were established as a means to provide the four necessary functions of scholarly communication to the academic community; namely registration, certification, awareness, and archiving in a single package (Prosser, 2003). This paradigm is often disadvantageous to the universities that produce the products published by these journals in several respects. First, the universities and research centers spend a large amount of funding in the form of departmental budgets and grants to produce some body of research. The author of this research must then submit it to a journal in order to have the necessary certification and publication performed and must in the process give up copyright to that work, requiring that author's host university library to invest further funding in purchasing license to that published work for use by others in the community (Gibbons, 2004, p. 13). This is tantamount to double-taxation on the university which must now pay not only to produce but also to consume its own knowledge. The University of California's Office of Scholarly Communication goes further to label the current scheme a crisis, and states "current scholarly publishing models are not economically sustainable. Researchers and students have access to a diminishing fraction of relevant scholarship" (The Regents of the University of California, 2009). Another way in which journal publication detracts from the institutional image is by the fact that journals publish research from all universities. This aggregation of diverse work from across academia can dilute a particular institution's contributions to a field of practice (Crow, 2002, p. 4), especially for younger, less well known research centers. A further limitation under the current system is the fact mentioned briefly in the section introduction that due to space and content restrictions, both the amount of research published in journals (i.e.

the number of articles) and the length (and therefore detail) of these reports is reduced (Smith, 2002, pp. 543-544; Genoni, 2004, p. 301). Add to this the length of time required by the formal process of publication in a journal and the inadequacy of the current model is readily apparent. The following paragraphs will analyze the journal services defined by SPARC in the context of institutional repositories and the manner in which these systems advance academia's ability to perform them.

Registration is the process of establishing the temporal priority of scholarly research, or in other words the official designation of publication dates for the purpose of establishing precedence among like research, traditionally recorded as the date a manuscript is received by a journal prior to publication (Van de Sompel, Payette, Erickson, Logoze, & Warner, 2004). Registration is therefore a very important part of maintaining the historical record of research, particularly in a distributed system of publication and communication. In the model given by ArXiv, the very act of submitting a manuscript to the digital repository fulfills the registration function (Van de Sompel, Payette, Erickson, Logoze, & Warner, 2004). This amounts to staking a claim, after which the certification process can begin using whatever means the community deems appropriate, be they the traditional methods of journal submission or the more radical approaches offered by institutional repositories (Crow, 2002, p. 13). The registration process is another way in which the incorporation of standards such as OAI-PMH is beneficial, as these standards ensure that the registration dates are recorded in a systematic manner to properly rank the temporal order of submitted works, providing the interoperability required to allow disaggregated repositories to function as a single system (Buede, 2000, p. 283). The benefits of this scheme are clear when one considers the limited forms of documentation that

are submitted to journals to begin with; under the model proffered by the advocates of institutional repositories, grey literature becomes a tool to establish the precedence of a scholar's work in an area by virtue of its residence in the IR, irrespective of its degree of readiness for final publication (Crow, 2002, p. 13).

The certification process is arguably the bulwark service offered by journals, and is likewise the greatest source of inertia in moving towards a new communication paradigm, since it is through peer-review that journals maintain credibility and respectability. Peer review is another area where the promulgators of institutional repositories differ, though less adamantly than with the issue of open access. Clifford Lynch, as a proponent for the coexistence of IRs and subscription journals, frowns upon the use of gate-keeping in institutional repositories inasmuch as such a function would deter users from contributing to the repository the sort of grey literature it was intended to host (Lynch, 2003). SPARC advocates a layered architecture, where the communications layer of the repository sits below a peer-review layer, the role of which can be performed by an "overlay journal" for the purpose of guaranteeing the content validity of the works in the repository (Prosser, 2003, p. 168; Crow, 2002, p. 13). As with all things the balance is likely in the middle. Lynch acknowledges that compartmentalizing an institutional repository system in such a way as to allow the community control over content in certain areas, as well as to produce some superstructure to enable gate-keeping mechanisms as in DSpace is acceptable, so long as these structures do not fundamentally alter the ability of the system to store less polished work in some way (Lynch, 2003), unlike subject repositories such as ArXiv.org which generally require that the papers submitted for inclusion be publication-ready or even referred in some way (Genoni, 2004, p. 301). In the 2003

Houghton Report designed to investigate the impact of Australian research organizations' adoption of IRs following the recommendations of a 2002 report by Australia's Commonwealth Department of Education, Science, and Training (DEST) information advisory committee, the authors recommended the adoption of standards to ensure the quality of IR contents in light of the "strong adherence to peer review." The report further recommended that a clear standard for identifying what quality control was present in a particular document be adopted, which amounts to ensuring that the metadata of a given document clearly specify its level of maturity and stage of review (Genoni, 2004, p. 303). This recommendation could be the middle ground the community is seeking, as it allows for clear identification of a document's certification without relegating the institutional repository to a proxy for the already well-developed journal publication system.

Awareness immediately stands out as one of the most beneficial aspects of IRs when compared with journals because an institutional repository increases the potential for exposure of a scholar's more current, un-published work. Additionally, one of the more problematic aspects of journal subscription is the high cost associated with them. With the proliferation of journals across many scholastic disciplines, it has become impossible for any library to maintain subscriptions to all the journals that would benefit its customers (Johnson, 2002, p. 5), resulting in a selection process which limits the research literature available to all the knowledge seekers within an organization. Institutional repositories remove the barrier of subscription cost to the public in viewing publication-quality work (Johnson, 2002, pp. 2,5). Furthermore, as concentrated archives of an institution's scholarly production, IRs reduce the diffuse manner of preservation of an institution's legacy and redirect the focus of the public on the university producing intel-

lectual capital rather than the journal containing it, building a powerful research history for the host organization (Crow, 2002, pp. 4-5). The converse is also true, however; namely that under the current scholarly communication system, publication in a respected journal bestows respect on the contributor (Smith, 2002, p. 543; Johnson, 2002, p. 2), a fact that produces one of the core elements of cultural inertia reducing the efficacy of the successful establishment of IRs, and one that the journal publishers are in no hurry to change (Johnson, 2002, p. 2). This implies the need to strike a balance between institutional disposition of research and publication in broader, more nationally-circulated media, a need which can be partly filled through commonality between systems such as the interoperability offered through standardized metadata formats which allow for harvesting by centralized search engines (Crow, 2002, p. 10; Shearer, 2003, p. 253).

The Archival power of an institutional repository is intuitively obvious given the discussion above, specifically when one considers the ability of individual researchers to self-archive knowledge that might not otherwise be published in any other way (Prosser, 2003, p. 168). In preserving these works in a centralized location, the researcher passes responsibility of maintaining these documents to the institution as a whole, which is better able to cope with evolving file format and other software changes, as well as the hardware technology changes (recall the opening discussion of 3.5" floppies) which drive obsolescence than could the author alone (Gibbons, 2004, p. 11). The rapid evolution of digital technology brings with it the danger that an institution may lose access to older digital documents (Heminger & Robertson, 1998, p. 159). When one considers the relative life spans of individuals and institutions, it is readily apparent that the longer-lived organization can accept a responsibility that no shorter-lived organism can to ensure

durable access to the output of that organism's efforts, especially when the information it seeks to steward is gathered in a central repository rather than dispersed across myriad local drives and individual web pages. With scholastic journals, archiving is limited to only those works that are accepted and published by the journal – all other knowledge produced by a scholar is consigned to oblivion unless preserved in some other way. Institutional repositories offer scholars a way to pass on the sum total of their legacy, and in the process enhance the legacy of the entire institution.

2.5 The Role of the Librarian

2.5.1 Librarians as Knowledge Brokers

As stewards of knowledge, libraries occupy a unique position in the knowledge chain. Rather than behave as simply a query tool, librarians use their insight and experience to connect knowledge-seekers with sources they themselves might not have expected, to include others seeking the same information, resident subject matter experts, or related resources to which the seekers might have been initially blind (Davenport & Prusak, 1998, p. 29). There is synergy in this relationship which is absent from even the smartest IT system due to the librarian's use of all five (maybe six?) senses when handling a request, which hallmarks some of the social aspects of Knowledge Management. At the heart of this exchange, and knowledge brokerage in general, is the process of *translation* (Davenport & Prusak, 1998, p. 98; Brown & Duguid, 1998, p. 103). Librarians liaise with faculty and understand the "needs and perceptions" of users (Bailey, *The Role of Reference Librarians in Institutional Repositories*, 2005, p. 266). Librarians work as boundary-spanners in both a market sense (connecting knowledge seekers with know-

ledge providers), and in the technical sense using their data gathering skills to translate a seeker's request for knowledge into a manuscript or other source that will satisfy that seeker's needs. In connecting individuals librarians fill the organizational role of "portal" within a university. In business, portals are managers who connect individuals with third parties capable of providing needed information (Hansen & von Oetinger, 2001, p. 114). This "matchmaking" role is informal in both business and university settings but provides obvious advantages to a knowledge seeker, and is powered by the loose disciplinary linkages between librarians and the communities they serve (Brown & Duguid, 1998, p. 103).

If "knowledge begins with an innovative mind that has connection to other minds" (Hoque, 2008), the librarian's function of making those connections provides a vital link in keeping the university a productive environment irrespective of the chosen technical solution. Moreover, a chosen technology should enhance, rather than interfere with the library's capacity to connect those minds to keep the institution a fertile ground for knowledge creation.

2.5.2 The Librarian's Role in Institutional Repositories

The major factors leading to the failure of an institution to populate an IR come from cultural, rather than technical concerns on the part of contributors. In his talk on institutional repositories and open access, David Prosser provides an anecdote about the effort to set up an IR in Minsk using the EPrints software developed by the university of Southampton: the major lead time was nine days to get the computer through customs and another 1.5 to set it up with the software (Prosser, 2003, p. 169). Underscoring that the technology is less of a limiting factor than other concerns, of the fourteen general and

specific factors listed for use or non-use of Cornell's institutional repository effort by Davis and Connolly, only two (learning curve and lack of functionality) were technical in nature (Davis & Connolly, 2007, pp. 13-16). The rest dealt with various concerns, but in the non-use category those concerns were primarily related to a lack of trust, in either the reputation of the repository as a quality source, or in the other users especially in regards to plagiarism or theft of intellectual property. Compounding these issues are the lack of a unified view of what IRs are and how they should be run, and clash between the drive towards open access and the protection of copyright (Genoni, 2004, p. 300). If one takes the Community of Practice view of IRs, it becomes readily apparent that member motivation is the key factor in ensuring a vibrant, active repository (Ardichvili, Page, & Wentling, 2003, p. 66). As knowledge brokers and information professionals, librarians can not only use their skills to help the organization determine the information needs of their organization (Henczel, 2000, p. 226), but can maintain their historical role as facilitators of collaboration and joint knowledge creation to encourage contribution to a lasting system.

The second aspect to maintaining vibrancy in any VCoP is to ensure the content is fresh, accessible, and as accurate as possible; and the library is in a prime position to ensure this happens in executing the role of maintainer of the repository. Since the IR represents the institution over the community, care must be taken to safeguard the reputation of the institution and thereby the perceived value of the content, which requires planning as well as oversight and management. The community of contributors is largely able to self-police when it comes to content validity, but when it comes to setting and enforcing standards of content and conduct, librarians as experts in collection selection

and management are ideally suited to ensuring the content is well-organized and well-maintained, thereby safeguarding the source-validity of objects within the repository. Their brokering abilities and lack of attachment to any particular discipline make librarians the natural choice for intermediaries should disputes arise, and in organizing and managing the repository they would have insight into its structure which would allow them to guide new consumers to and through the repository in much the same way as they instruct knowledge seekers on how to use extant search technology.

In order to cope with the evolution of information technology and the risk of format obsolescence, the university will likely need to designate an entity to specialize in the digital preservation of its legacy (Heminger & Robertson, 1998). As with collaboration, the library's expertise in this role has been honed over millennia making the library the natural organization to care for the university's knowledge base, in concert with the university's computer services and IT professionals (Gibbons, 2004, p. 17). The centralization offered by institutional repositories removes the discovery and selection concerns of the arm of the library charged with keeping the university record current and available, thus easing the task of preservation.

2.5.2.1 Communication and Publication

Under the current model of scholarly communication, Librarians as stewards of collections support the awareness and archiving functions by selecting journals to include within the library collection and maintaining those journals in perpetuity. With IRs, the mantle of exciting awareness falls squarely on the shoulders of librarians, as it becomes their responsibility to ensure that the material within the repositories is properly tagged and accessible to (and thus interoperable with) the community at large

(Crow, 2002, p. 9). The archiving role expands as well, as libraries and their sponsoring organizations become singularly responsible for stewarding the intellectual wealth of the institution.

2.5.2.2 Collection Management

The principles of library collection development and management have evolved from their first codification in the 1970s (Allard, Mack, & Feltner-Reichert, 2005, p. 325) to include the incorporation of content, irrespective of medium, “in whatever form it is acquired or provided by a library” (Genoni, 2004, p. 303), and thus remains well within the purview of librarians. With the advent of self-archiving authors as major selectors of content, the acquisition function of the librarian moves largely to the community, provoking some shift in the librarian’s responsibilities away from acquisition towards stewardship (Allard, Mack, & Feltner-Reichert, 2005, pp. 325-326). The librarian is not, however, completely divested of the role of selector; indeed, if the DAEDALUS experience produced any lessons, one of the more important was that librarians are instrumental in targeting and encouraging the upload of relevant material, especially during the initial growing pains of institutional repositories (Mackie, 2004, pp. 3-4). Furthermore, librarians as brokers possess a body of broad institutional memory developed over years of interacting across disciplines. This memory is invaluable in identifying areas of research that may not otherwise be targeted for inclusion in the repository (Bell, Foster, & Gibbons, 2005, p. 288). Building a collection requires expertise in selection, description, cataloging, storage, and management (Genoni, 2004, p. 303; Gibbons, 2004, p. 8). Combating the malaise and inertia impeding faculty contribution requires education, advocacy, active recruitment, and broad knowledge of the intellectual property limita-

tions to populating IRs with published materials, functions which a dedicated staff conversant in these issues is far better prepared to grapple with than the otherwise-engaged faculty.

Librarians as information professionals are in the correct position to identify the needs of their host institution and its legacy and use that analysis to target content for use in the repository (Henczel, 2000, pp. 214-215). An *information audit* is a formal process that determines what information needs the members of an organization have and how those services are used, enabling the organization to complete a larger *knowledge audit* to define the institution's knowledge assets, which should be the first step in implementing any knowledge management system (Hylton, 2002, p. 2). The knowledge audit often includes business and cultural assessments, which help to define the overall climate of the organization, and an analysis of where the knowledge gaps are, in what areas knowledge is best applied, what is accessible, and what is contained by the organization (Liebowitz, Rubenstein-Montano, McCaw, & Buchwalter, 2000, p. 5). This ongoing analysis of institutional needs is an excellent place for librarians to put to use their professional skills in evaluating the university's intellectual and cultural status to help develop an overarching strategy for managing a knowledge repository (Henczel, 2000, p. 226).

One area in which most of the IR magnates agree is the vital importance of ensuring that the files reposing in the IR have accurate, properly formatted metadata (Gibbons, 2004, p. 7). Properly formatted metadata (or data about data) using an open standard is the engine that makes institutional repository content available to the entire scholastic community by ensuring interoperability (Crow, 2002, p. 10) and mitigation of

the troubles wrought by dead links, so much so that D-Space incorporates special handles as a mandatory part of the metadata to ensure accessibility (Smith, 2002, p. 547). As catalogers and bibliographers, research librarians are by definition experts in metadata and the way in which the application of key words and other tags affect the ability of a document to appear in a subject area search. Automated IR submission forms require that certain field such as author name and paper title be entered prior to the system's accepting the document. The development of additional standard metadata fields to tailor individual collections, along with auditing submissions to ensure the adequate definition of metadata are functions well within the area of expertise of librarians (Allard, Mack, & Feltner-Reichert, 2005, p. 334), and form a part of the continually-evolving role of the reference librarian.

2.5.2.3 Other Roles

The library has several assets that place it in an excellent position to manage the institutional repository. As the library assumes this responsibility, there is synergy between the library's skills and the roles required to manage the IR that will become important to its maintenance, ensuring continued contribution, and encouraging use. The library is usually seen as a public good, much like the knowledge it stewards (Ardichvili, Page, & Wentling, 2003, p. 69), and this sets the library apart from interdepartmental squabbling. The library's first and potentially greatest asset is the level of trust it has within the community of the university, and the large network of interrelationships librarians have as a bridge within and between organizations (Gibbons, 2004, p. 17). These relationships are very useful in lending credibility to an IR undertaking, which catapults the library into the role of advocate and arbiter of change (Bailey, The Role of

Reference Librarians in Institutional Repositories, 2005, p. 266). This trust is also beneficial when coupled with the notion of legacy in recruiting works from generative faculty who are eager to leave a lingering bequest to the university (Bell, Foster, & Gibbons, 2005, p. 289).

Recruitment and combating culture are in general very important facets of ensuring the success of the IR, and librarians as information professionals have several skills which greatly improve their ability to perform this function. Information audits and institutional memory have been discussed already, and contribute to librarians' ability to identify potential works for inclusion in the IR. Librarians can use their skills in collection definition to provide guidance in developing digital collections (Allard, Mack, & Feltner-Reichert, 2005, p. 333). Librarians can also use resources such as RoMEO and other services offered by SHERPA to develop a list of publications that have copyright policies which are friendly to publication in the IR (Bell, Foster, & Gibbons, 2005, pp. 288-289). As Morag Mackie discovered through this process, most academics have little issue with a third party identifying which of their published works held friendly copyrights so long as they weren't bothered with the details, and further, by adopting a policy which assumed tacit approval on the part of the faculty to post this literature unless a faculty member specifically requested that a work not be included allowed the library of the University of Glasgow to populate the IR with journal papers far more rapidly than would have been otherwise possible had they requested permission for each submission. This also speaks to another form of inertia, namely that authors bogged down with other work are less willing to commit to self-archival, especially with a new, unproven system (Mackie, 2004, pp. 2-3). By taking these initial steps to recruit content, the library is

able to jump-start the IR population process and help the repository to build credibility. Once the low-hanging fruit has been picked, the library can turn its full attention to the active recruitment of grey literature, which is also vital in establishing the intellectual legacy of the university. According to Peter Lyman and Hal Varian, more than 55% of hard drives storing original works are local (Lyman & Varian, 2000), which is a medium that is not easily accessed and sifted through after its time of regular use ends. In the words of Susan Gibbons: “doing nothing to preserve digital works of enduring value guarantees their loss” (Gibbons, 2004, p. 11), which infers that the library must make recruitment of grey literature a priority, since it is that literature which has the least chance of being preserved otherwise. A strategy that quickly produces a sizeable population of useful literature will help to motivate further contributions. System success begets broader use (Bell, Foster, & Gibbons, 2005, p. 289), and so the recruiting activity is vital in ensuring the use of the IR and thereby its success.

Education and training of both authors and knowledge seekers is another necessary element of the proper use of a system. Librarians already train new students in the various search tools available to them to locate knowledge, so the expanding training to embrace another new system is a natural progression in the role of the librarian (Bailey, *The Role of Reference Librarians in Institutional Repositories*, 2005, p. 266). Training brings with it the added benefit of increasing exposure to the burgeoning system, which acts as its own form of marketing, and can include some of the nuances of copyright and other limiters to help faculty understand the full limitations presented under the current scholarly communications paradigm and the advantages offered to both the individual *and* the institution by the IR. Even after the completion of formal training, the library can

remain involved in assisting users and walking faculty through the submission process and provide guidance in entering metadata (Allard, Mack, & Feltner-Reichert, 2005, p. 334). They can also help develop the interface of the system and use feedback from the training process to recommend adjustments to make the system as easy to use as possible, thus further removing barriers to its embrace by the institution.

These roles are but a few of the possible ways in which librarians can expand their current role to better serve their customers in managing digital information repositories. Understanding software, project planning and management, collection definition, metadata guidance, submission review, author training are all roles that are not at all dissimilar from those already performed by research librarians (Allard, Mack, & Feltner-Reichert, 2005, pp. 333-334), and consistent with the theme of this section, represent a natural evolution of the library professional with the advent of new technology and new ways of doing business, within the context of knowledge management. The next chapter will propose a method to further develop these ideas and apply them to developing a framework within which to answer the questions posed in the introduction.



3 METHODOLOGY

Having reviewed the evolution of library knowledge stewardship and generation, the study now turns to present and future knowledge management initiatives underway in the academic and research arenas to begin to develop an understanding of the current state of the practice and how well some of these fledgling systems are fulfilling the roles for which they were envisaged. In particular, much of the material pertaining to institutional repositories was written during and shortly after the initial establishment of IRs at various organizations, in the period from 2002-2005. To better guide the AFIT library in determining an institutional repository approach to suit its particular needs, it is appropriate to gather more current reviews of some of these systems and some of the relevant reasons for their successes or failures. In addition, some direct purposive canvassing of several libraries may provide insight which could aid the AFIT library in analyzing its needs and developing its own institutional repository or another KMS.

The literature review for this investigation is structured to provide a summary of the recent published works that encompass the major themes of this study (knowledge management, library history and function, and institutional repositories). This review provides the basis for developing a list of questions for use in a semi-structured interview, which is designed to acquire more current answers to the research questions posed in the introduction along with the context relevant to the respondents.

3.1 The Case Study Method

The method employed for this research was a case study of single case design with multiple units of analysis. An exploratory case study method is appropriate for this investigation in that it attempts to ascertain *what* libraries are doing, *what* they have planned, and *what* they find to be important (Yin, 2003, p. 6). The particular data collection methods described in §3.2 below, namely interviewing and analysis of items of public record, support the discovery of prevalence within the cases (Yin, 2003, p. 7), which is in keeping with the desire to obtain descriptive information about IT trends that can be used to help guide the efforts of a library seeking to implement a particular class of system by providing evidence about what other libraries are doing and their experiences with relevant technology.

3.2 Data Collection

Data collection followed a relatively simple procedure. Source institutions were identified through the procedures described in the next section. Once a list of potential sources was gathered, particular respondents were selected out of the respective institutions' organizational structures based on position description, or through directly contacting the heads of relevant divisions and asking them for direction to the best respondent. Once individual respondents were identified, they were contacted directly (via telephone or e-mail) and asked to participate in an interview, either over the phone or in person (based on location), and given an overview of the research purpose and interview format. Willing respondents were offered a copy of the questionnaire (see appendix 1) in advance to allow them to gather any statistics they might not have handy and asked for

an appointment time during which they would discuss their responses to the questionnaires and expound on any as need be. Data was input directly onto a copy of the MS Word™ document and saved with the institution's name for later analysis.

3.2.1 The Interview

Interviewing provides several advantages over direct surveys in that it allows for gathering of historical information (Creswell, 2003, p. 186), collection of nonverbal/unwritten behavioral data (Schwab, 2005, p. 41), and interviewer intervention and clarification where necessary and appropriate (Neuman, 2006, p. 306). This method involved largely telephone interviewing accompanied with opportunity- (i.e. location-) based face-to-face interviews where possible. The method enables timely gathering of information from geographically-dispersed respondents and largely reduces the potential for non-response since positive contact and appointments directly ensure that the respondent does indeed have time for the interview and will provide answers to it. By keeping the structure loose and allowing for probing questions and tangential discussion, a semi-structured interview allows for expansion on themes that may not have occurred to the interviewer when developing questions, and may expose the interviewer to additional sources of evidence which better explain the phenomena being examined and which help provide a richer, more holistic view (Yin, 2003, p. 90; Patton, 2002, p. 339).

In keeping with the philosophy of holism, responses to a semi-structured interview have the added benefit of providing answers to questions of both closed- and open-formats, both of which are designed to allow full-responses to reduce the likelihood of misinterpretation of the question or the response (Neuman, 2006, p. 307). Qualitative interviews are a good source for positivist (that knowledge based on experience) and interpretive

data (Meyers & Newman, 2006, pp. 2-26), which supports the research goal of determining the state of the practice and collecting data that will provide some normative heuristics for the purpose of guiding the AFIT/AFRL Library's own initiatives.

3.2.2 Ancillary Data

In addition to the data collected through interviews, this study utilized items of public record to provide support and context for interview responses. This documentation took the form of the official web sites of those institutions interviewed and those institutions and systems identified by respondents during data collection, and is presented in the analysis portion of this report where applicable to give further insight into the responses of the interviewed organizations. Additionally, documentation of this type provides a stable, unobtrusive, exact, and broad resource for use in analyzing interview responses (Yin, 2003, p. 86). Additionally for the purposes of this particular investigation, access (or lack thereof), normally considered a weakness, relates directly to the open source discussion in the literature review, and can provide additional fodder for discussion.

3.3 Sample Selection

In keeping with the principles of representativeness, heterogeneity, and specific comparison (Maxwell, 2004, pp. 87-90), the libraries chosen for the interview were selected to obtain a cross-section of US academic libraries, based on size, geography, focus (education, research, or some combination thereof), and the stated purpose of the library. Special care was taken to interview personnel at those libraries revealed by the literature review and other sources as having been involved in the development or institution of

institutional repositories, since that particular knowledge management system is under direct consideration by the AFIT/AFRL library for implementation. Target institutions included those that have undertaken an IR-type effort for the purpose of stewarding institutional produce, such as the University of California; government academic libraries such as the Defense Language Institute's (DLI's) Aiso Library and the Naval Postgraduate School Library to provide relevant context for the AFIT/AFRL library's own plans and initiative; and smaller academic libraries for comparison/contrast with broader regional centers. Though the sample was targeted to show variety in location and size, samples were not randomly selected. However, some randomness is incorporated through snowball sampling, where interviewees are asked to propose additional subjects for questioning to expand the subject pool to encompass institutions that may not have otherwise been directly considered by this study.

To ensure that subjects respond from a position of authority, individual respondents at these institutions were selected on the basis of their proximity to the level where policy is instituted (i.e. a policy maker or designee).

3.4 Data Analysis

The main method of reducing the collected responses to the questionnaires was through categorization, or the coding of responses to identify larger themes (Maxwell, 2004, p. 95). Participating institutions were first classified according to size, primary purpose, and affiliation (in this case public, private/nonprofit, or DoD) to allow for examination of differences in responses. The questionnaire was then decomposed into its

major theme areas and the responses of the participating organizations analyzed and sorted within those themes according to category.

3.4.1 Content Analysis

The final stage of this method was a content analysis of the data collected, a technique for drawing inferences from texts to their use contexts (Krippendorff, 2004, p. 18). Responses were summarized in tables (located in Appendix II) by question area and institution which provides the ability to discover patterns and relationships that might be otherwise overlooked (Krippendorff, 2004). Responses were then compared with archival data collected from public records described in §3.2.2.

3.4.2 Mapping Methodology Elements to Research Questions

The below table demonstrates the manner in which the methodology models the research questions. The numbering defines the degree to which each method provides information about the relevant topic area as posed in the research questions, with “3” defining a significant contribution and “0” no contribution.

Method	Past/Current Trends	Planned Activities	Useful Unplanned Activities
<i>Literature Review</i>	3	0	2
<i>Public Source Investigations (Institution Web Sites)</i>	3	1	2
<i>Interviews</i>	3	3	3
<i>Analysis</i>	3	3	3

Table 3-1: Mapping research methods to library activities

3.5 Validity

3.5.1 Construct Validity

Construct validity determines how well the measures of a subject represent the concept for which the model was developed (Schwab, 2005, p. 16). In this case the purpose of the research was to examine the IT initiatives of various academic and educational institutions to develop a snapshot of how some of these technologies are being used in practice and to provide the AFIT library some insight into the experiences of these institutions for use in planning their own initiatives. The research design accommodates the first portion of this by selecting a broad range of institutional sizes and specialties, and approaches the second portion of this purpose by identifying organizations that share some of these characteristics with AFIT which in turn makes recommendations more relevant. Further, in keeping with the validating use of multiple sources of evidence (Yin, 2003, p. 34), the methodology combines interview data with information garnered through items of public record, namely institutional web sites to provide some context and corroboration for responses to interview questions.

3.5.2 Internal Validity

Internal validity describes the degree to which “variation in scores on a measure of an independent variable is responsible for variation in scores on a measure of a dependant variable” (Schwab, 2005, p. 14). As a quick-look, descriptive study, this research design does not seek to draw causal claims from the research questions but rather to examine the current state of the practice and update the literature, thus internal validity is not applicable.

3.5.3 External Validity

External Validity describes the generalizability of data and exists when findings of a study are generalized appropriately (Schwab, 2005, p. 19). External validity is determined by the research design; for the purposes of this study the sample was selected in such a way as to provide breadth of library size and purpose to allow for discrimination in responses based on these differences, and thus to determine whether and how the needs of the various institutions and their responses to those needs differ based on these factors. Snowball sampling helped to ensure the identification of institutions not in the initial pool of prospective subjects to help remove researcher bias in subject selection and thus allow for a more general view. Additionally, the instrument was shared with two experts in the field, one of whom was not an interview subject, in order to gain an unbiased appraisal of its thoroughness. In both cases, the script was deemed comprehensive and an adequate method of assessing the status of the subjects' IT efforts.

3.5.4 Reliability

Reliability describes the replicability of data; specifically the idea that data collected across multiple collections are consistent, whether those collections be separated by measurement items (internal consistency), raters (inter-rater reliability), or time (stability) (Schwab, 2005, p. 241). Reliability is governed by the data collection procedure and is in this case supported through the standardized interview procedure detailed in §3.2.

3.6 Summary

This methodology defines an approach designed to provide rich data through a set of open interview questions and probes as presented to policy-makers or designees at several key institutions, and content analysis of narratives returned from these interviews. The interview method allows for detailed responses to a battery of questions designed to assess the current state of library KM and IT initiatives across a broad range of academic and research institutions.

The results of this method will be largely descriptive due to the free form of the interview process, holistic due to the interrelatedness of KM and culture, and will draw some conclusions from literature and collected data to help guide academic libraries today in their effort to implement technological systems to serve their customers.

4 RESULTS

This purpose of this study was to procure a snapshot of the current efforts of various academic and research libraries to incorporate some of the new technologies described in the previous chapters, and thereby to start to form a practical appreciation of some of the real-world challenges to standing up systems like institutional repositories. The study attempted to identify institutions of various sizes, data needs, purposes, and affiliations to gain insight into how these differences in purpose manifest in different desires and applications of the available information technology (IT) systems.

4.1 The Sample

A structured interview method was selected to gain this insight, a process which involved selecting organizations to participate, contacting librarians or IT professionals within these institutions identified by their respective organizational structures, setting up appointments with those individuals or their designees, and conducting the interviews either over the telephone or face-to-face, based on the distance between the interviewer and respondent. The sample was selected through a combination of opportunity and snowball sampling. The first round of institutions was selected based on their stated purpose and proximity to the interviewer; the second through referral by respondents in the first round.

The specific academic and research centers involved in this study were The University of California, Berkeley (UCB), the Naval Postgraduate School (NPS), the Air

Force Institute of Technology (AFIT), Wright-State University (WSU), the Defense Language Institute (DLI), the Monterey Bay Aquarium Research Institute (MBARI), the Patriarch Athenagoras Orthodox Institute (PAOI), and OhioLINK.

4.2 The Questionnaire

The questionnaire comprised a 98-question, hierarchical form. The first tier was composed of 36 questions mapped directly to the research questions, with additional sub-questions designed to provide further resolution in answering those top-tier questions. Questions were developed to target themes identified in the literature review. The first section of the questionnaire was designed for use in classifying the library in question, asking for demographic information such as the size of the institution's community, distribution of holdings, targeted customers, and so forth.

The second set of questions was geared towards assessing the current activities of the respected libraries, with emphasis on systems like institutional repositories (IRs) or other infrastructure designed to store and share intellectual products. Questions were organized to prompt discussion about the various systems in use in the library, their purpose, administration, barriers to their establishment, and the effect of their implementation on the staff and institution. The third set of questions targeted upcoming, planned activities by those institutions to develop various IT systems, and largely restated the questions of the second section in the future tense. Both sets attempted to gain insight into the institutions' technological and cultural planning efforts, as well as their criteria for defining system success.

The final portion of the questionnaire was a more free-form section, designed to identify other technologies librarians felt might be valuable in serving their customers that may not have been formally identified in any planning sessions, as well as any potential barriers to the implementation of those systems, and a final question to allow respondents to discuss anything else that was not covered explicitly by the questionnaire. The purpose of this section was to provide the interviewer an opportunity to gain insight which might be otherwise limited by the specificity of the other questions and to help identify potential areas for future study.

4.3 Results

This section summarizes the information gathered from the interview process. The results are presented here in a largely narrative form, with thematic summary tables of interview responses located in Appendix II. These tables were created by distributing condensed responses to the questionnaires by institution against the top-tier questions and sorting to identify commonalities. The presentation follows the overall layout of the questionnaire itself, though responses that were given in later areas of the questionnaire may be presented earlier where relevant for comparison with responses from other institutions. This is due to the fact that some amount of redundancy was built into the questionnaire to ensure the capture of information that might not have occurred to a respondent along a particular line of questioning, or that was recalled later during the interview. As discussed in §4.2, questions 2 and 3 contained some sections which differ only in regard to tense. These responses are also condensed for presentation purposes.

4.3.1 Library Classification

As stated above, this study was conducted on eight academic or research-oriented institutions, identified through a combination of purposive and snowball sampling. These organizations are categorized below according to population served (or rather the organization's target population, since web-based services are often accessible to the entire world), their jurisdiction (public, private, or defense), and their focus (research, education, or some combination thereof). One of these eight, OhioLINK, is a statewide consortium of various public and private institutions and is included in this study for its deep involvement in institutional repository, archival, and data sharing initiatives. Its service population is an order of magnitude larger than that of the next-largest single campus and so represents an outlier for certain purposes, however its inclusion will become readily apparent as it provides insight into several themes that will be discussed below.

Direct categorization of the remaining seven institutions is distributed in the following ways. For population served, the institutions can be grouped into small (populations less than 1000), medium (populations between 1000 and 10,000), and large (populations greater than 10,000). Jurisdictional groupings include Public, DoD, and Private Non-Profit. Finally, focus groupings include research, education, and a combination of both. These categorizations are summarized in table 4-1 below.

Library	Population	Jurisdiction	Focus
OhioLINK	600,000	Public	Communication
UC Berkeley Library	37,437	Public	Research/Education
Wright State University Library	16,000	Public	Research/Education
DLI Library	4,500	DoD	Education
NPS Library	2,500	DoD	Research/Education
PAOI Library	1,450	Non-Profit	Education

AFIT Library	750	DoD	Research/Education
MBARI	500	Non-Profit	Research

Table 4-1: Institution Categorization

4.3.1.1 Customer Distribution and Resource Location

Of the institutions questioned, those that maintained a distinction between local and remote users tended to focus primarily on local use; with NPS maintaining the largest separation (75%/25%) between local and distance learners. The other institutions were in the 90-100% local range as far as focus. The larger organizations did not maintain statistics in this area, citing a difficulty in doing so; for example Berkeley defines users in the residence halls as local and thus does distinguish between website hits from within and outside of the library. OhioLINK, as a gestalt repository of repositories, defines its population as entirely remote and seeks to provide content to the widest range of potential users over the web.

As for resource location, all of the institutions questioned used some form of html-based retrieval system, both for local and remote users. These systems tended to take the form of an online catalog with some form of web interface, with the exception of the PAOI which used a relational database called FileMaker that was not accessible outside of the local area network. However, with the migration of its collection to the larger Graduate Theological Union's (GTU) library, the PAOI's works are now available on that institution's online catalog and the WorldCat, a resource location system that allows for the searching of books and other media located in thousands of member libraries around the world (OCLC Online Computer Library Center, Inc., 2009). MBARI does not make its library print resources available outside of its LAN, though it does make a

substantial amount of other data and media freely available through its multi-faceted website <http://www.mbari.org/>.

4.3.1.2 Library Use

All of the libraries canvassed in this investigation cited multiple local uses of the library facility by daily patrons. The educationally-focused (and research/education-focused) institutions provided patrons space for study and relaxation, collaboration, software application use (i.e. the MS Office suite), and other services (such as photocopying), in addition to source location; the singular research-focused institution emphasized source location as its primary purpose. Wired internet access was a consistent theme across these schools as well, with Berkeley, WSU, and DLI offering some form of wireless internet access as well.

The most frequently cited remote use for library patrons across research oriented institutions was some form of proxy authentication for the purpose of access to subscription journals. This included OhioLINK which provides students, faculty, and staff of its member institutions authenticated access to journals for which they are licensed, especially those contained within its member databases. The two education-focused institutions noted catalog and database access as common remote uses.

4.3.1.3 Library Holdings and Circulation

All of the libraries canvassed held a variety of content, especially books, journals, and locally-produced works. The nature of these local works differed among the institutions, with research/education-focused organizations housing their graduate students' theses and dissertations, while DLI's local content was more geared towards its students'

course materials and curricula, PAOI's with its lecture series, and MBARI's tech reports and media.

As far as circulated items, books were the most popular media loaned across all respondents. AFIT, NPS, DLI, and Berkeley also cited multimedia (i.e. audio and video) loans, though these tended to be underutilized in comparison with print media with the exception of DLI which focuses on language instruction. In the case of MBARI, the library's direct jurisdiction is largely over print media, with most other forms of multimedia available over the internet through the website.

4.3.2 Library IT Initiatives

4.3.2.1 Library IT Use and Origins

All of the organizations represented in this study have used IT in some form or another since the early nineteen-nineties, with Berkeley, Wright State, the Naval Postgraduate School, and AFIT as far back as the early eighties.

Most of the libraries interviewed use some form of integrated library system (ILS), which combines catalog search and retrieval with loaning and other collection management software. UC Berkeley and the PAOI used home-grown systems, while the other centers used some form of commercial application. Other common applications were user software and internet/e-mail access, backend support software, proxy or virtual private networking (VPN) services for accessing subscription media, and catalog services. In judging the popularity of applications, many of the institutions had no real metrics for local use (insofar as there aren't hit-counters or other measurement devices on applications installed on local terminals). Some of the mentioned local uses were internet and office applications, the catalog, and MBARI mentioned ZOTERO, a plug-in

for Firefox designed to capture article citations. Remote use included aforementioned VPN and proxy services, along with catalog and other resource systems for the institution.

4.3.2.2 External Linkages

Responses to the interview questionnaire described external linkage of the physical collection in two different ways. Linkage of the collection to a parent- or sister-organization was interpreted as physical sharing of resources or some hierarchical ownership of the collection. The AFIT collection is shared with AFRL, the PAOI collection with the Graduate Theological Union, and some portion of MBARI's collection with the Moss Landing Marine lab. The second view of linkage is one of external access to browse the catalog. Six of the eight institutions mentioned the Online Computer Library Center (OCLC) explicitly – that either they were already a member or were transitioning to membership. OCLC is a “nonprofit, membership, computer library service and research organization dedicated to the public purposes of furthering access to the world's information and reducing the rate of rise of library costs” (OCLC Online Computer Library Center, Inc., 2008), which links the collections of more than 69,000 libraries around the world. The education (and research/education) institutions also had some formal interlibrary loan agreements set up with other institutions, most of them via electronic means.

4.3.2.3 Intellectual Production and Stewardship

As stated in §4.3.1.3, the research/education institutions all produce intellectual work such as theses, dissertations, and journal articles. MBARI produces journal articles and various inventions and patents, and also receives theses and dissertations through its

collaboration with the Moss Landing Marina Lab. The education-focused institutions produce curricula and other forms of intellectual output such as PAOI's lecture series, and with its full membership in the GTU, PAOI will also begin to produce graduate theses and dissertations.

The question of stewardship of these works introduced the topic of institutional repositories for the first time in each interview session. All of the canvassed institutions were aware of the emerging technology either through direct use or through conferences, and all of the research oriented organizations either used or planned to use some form of IR for a portion of its output. DLI's library maintains an institutional archive of the history of the institution, but that archive is fairly static and not well characterized. For those libraries using IRs, the method of use was split between a regional or subject-based system (Berkeley, AFIT), and a hybrid regional and local, institution-hosted system (in the case of NPS, MBARI, and WSU). Wright-State University plans for both a locally-hosted system *and* continued use of the larger system hosted by OhioLINK, and the Naval Postgraduate School continues to populate the Defense Technical Information Center (DTIC) while developing its own repository system. MBARI provides content to the International Association of Aquatic and Marine Science Libraries and Information Centers (IAMSLIC), which hosts a subject repository called the Aquatic Commons <http://aquacomm.fcla.edu/>, but also maintains an extensive data-sharing website for direct popular consumption. AFIT is gearing its efforts towards developing a seamless interface with DTIC to enhance the experience of its customers, and UC Berkeley has refrained from creating its own repository system due to the existence of the California Digital Library (CDL, described in Chapter 2), a unified digital repository for the entire

University of California. DTIC, the CDL, and OhioLINK all represent large, state or nationwide repositories designed to capture the intellectual output of a host of institutions, and provide certain other features such as federated searching and large catalogs for all of their member institutions. OhioLINK in particular is designed to provide seamless access to its holdings through member institutions' websites and employs D-Space architecture with much of the functionality described in chapter 2 to do so.

In recognizing the need to develop systems to better serve their patrons, the research (and research/education) libraries all employed some sort of active means of interviewing their customers, either through surveys or interviews with customers, focus groups, or other informal approaches such as MBARI's all-hands brown-bag lunch meetings. In addition to these, Berkeley, WSU, and DLI directly mentioned conferences in response to these questions, with NPS, AFIT, and MBARI inferring formal and informal conferences as methods of identifying new technology. OhioLINK cited its origins as an initiative following a recommendation by the Ohio Board of Regents to create a statewide electronic catalog system. OhioLINK's website goes further to state that the motivation to continue to expand services comes directly from Ohio's desire to continue to pioneer library automation (The Ohio Library and Information Network, 2008), in keeping with earlier advances such as the development of OCLC (now the Online Computer Library Center, previously called the Ohio College Library Center). PAOI's need to develop its IT systems was recognized directly by a need to manage a burgeoning collection, and the transition to GTU was recognized due to the limitations of their physical facility to safeguard its rare manuscripts (i.e. inadequate fire suppression and security

systems) as well as the limits of the IT system to provide those works with the necessary outreach to consumers.

Configuration of IR systems among the research libraries (and OhioLINK) varied widely among the sample. Five of the respondents mentioned some sort of configured, off-the-shelf system to manage their IR (or the partnered system they accessed). AFIT's plans are to coordinate a homegrown interface with DTIC to manage its repository function. D-Space was mentioned by three of the respondents, namely OhioLINK which uses it as its primary infrastructure, WSU through its collaboration with OhioLINK, and NPS as an alternative. Berkeley lacks its own repository system, but contributes to the California Digital Library which uses the UC's eScholarship software, and the aquatic commons is powered by the University of Southampton's EPrints 3 suite (Aquatic Commons, 2008).

In regard to available content, the education/research centers all listed theses and dissertations among the works they steward in their repositories. MBARI noted various scholarly materials; both published and unpublished as reposing at the aquatic commons. The military research/educational institutions included various tech reports, departmental reports, and other projects and publications. As far as efforts to expand collection, NPS and WSU mentioned a desire to steward faculty grey literature explicitly, and while Berkeley did not mention directly, their section of the California Digital Library (CDL) is replete with working papers and other grey literature from students and faculty, as well as finished open journal articles (The Regents of the University of California, 2008). AFIT expressed the desire to include CV citations, and NPS to add other course-oriented content such as slides. The two large institutions and OhioLINK

also noted the preservation of institutional legacy and other archival functions as falling within the purview of the repository, a view supported by the CDL as well.

Responses to the role and responsibility distribution in the IRs varied; though generally the targeted consumers were listed as “all” or “everyone,” and contributors included faculty and students almost universally. OhioLINK defined its contributors in terms of who loads content, namely the university liaisons, though the material is up to their discretion and thus can be inferred to be faculty and students, and IAMSILIC accepts reports from its member institutions – whose faculty (i.e. researchers) produces them. All five research institutions described content handling as being under the purview of the respective library, with system management falling under the owning systems division, especially in cases of collaborative systems. The institutions’ approaches to managing submissions generally involved some form of mediation. WSU and NPS expressed a desire to move to more direct loading by content providers, though with some system in place to ensure adequate population of metadata and marking information. OhioLINK’s content is generally loaded by the liaisons directly, except for some bulk submissions which OhioLINK will load off of a mailed-in CD or some other medium on behalf of an organization.

Quality control, metadata, and interaction, were all described previously as very important aspects of a well-run IR-type system, and all of the institutions questioned who had or were planning IRs spoke of metadata and ensuring its proper formatting to some degree. OhioLINK, WSU, and NPS mentioned standard OAI-PMH compliance like Dublin-Core and MARC (in NPS’s case this compliance was planned for the future system, the current system is not compliant). Further, according to its website, “DTIC is

a registered Data Provider with the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) Registry” (The Defense Technical Information Center, 2008), thus any seamless interaction between AFIT and DTIC will also be OAI-PMH compliant. The EPrints package used by the Aquatic Commons touts itself as being “the first professional software platform for building high-quality OAI-compliant repositories” (EPrints, 2009), and the Aquatic Commons sites itself as complying with OAI 2.0 standard (Aquatic Commons, 2008). The California Digital Library also uses OAI architecture (The Regents of the University of California, 2009), thus all of the institutions interviewed that are involved with repositories use an OAI-compliant structure. Some limitations mentioned by OhioLINK and MBARI with regard to metadata were the fact that the popular Dublin Core metadata standard is general and does not provide enough standard fields for particular areas of science such as phylogeny or marine biology which have additional categorization schema not covered by the standard library metadata structure. To help mitigate this, OhioLINK provides additional standards (such as Darwin core which adds categorization structures such as kingdom and phylum for species) for specific collections which can then be accessed by non-OAI harvesters. The CDL cites this practice as well, stating that “data providers also decide which metadata formats to expose for harvesting, beyond the one required data format of simple Dublin Core” (The Regents of the University of California, 2009). Interaction on all these OAI-compliant systems is as described in the literature review portion of this study: through metadata harvesters such as OAIster or Google Scholar.

As far as quality control, all of the respondents mentioned some aspect of the submission process as a primary gate keeping function. OhioLINK relies on the member

institutional liaisons to control quality according to their institutional standards. WSU employs staff mediation at entry – a metadata librarian populates the metadata and the documents are audited for completeness, and likewise under its current system NPS has a single point for data entry. DTIC's own submissions process controls metadata on its site for submitting institutions such as AFIT. MBARI's data sets are automatically tagged through XML at the time of collection, and the IAMSLIC board approves and does quality control for submissions to its repository. NPS plans to incorporate a workflow scheme under its new system, where faculty can directly submit packages via a form which will require certain tags, but that submission will then be examined by staff for completeness prior to posting on the system. Something similar was discussed with OhioLINK in relation to the University of Miami's move towards author self-publication. Here again the liaison would be the key to ensuring that the quality standards of the system were met prior to final posting.

In judging whether or not IRs were successful, four of the six institutions involved in IRs mentioned hits and click-through statistics as indicators of use which go towards the definition of success criteria. Google analytics was a commonly-mentioned tool for determining the use of these systems. From among the institutions that already had active IRs at the time of the interview, OhioLINK and MBARI (speaking from its perspective of the aquatic commons) defined theirs as successful. OhioLINK's criteria for that evaluation were that the system as it exists fulfills the functions promised to the members at the time it was initiated, though it is by no means a finished product and there is a continued effort to improve it and provide the best content possible. MBARI cited the deposit rate of documents and the tracking hits as indicative of the success of

their linked repository, and also noted plans to improve on and expand its reach through the incorporation of additional languages. NPS considered its current IR-like system to be a marginal success – it succeeds in stewarding works, but is not always available and is incredibly difficult to use and keep current – and for that reason is embarking on an effort to stand up a modern repository to satisfy the current system’s deficiencies. WSU’s efforts are still recent (their repository was started up in October 2008), and so the database is not well populated or used at present. Other criteria for success mentioned by the polled institutions were the ability of the system to seamlessly deliver the desired content to the knowledge seeker, ease of use by contributors and managers, and availability of stewarded works.

All of the polled institutions described building advocacy as key to ensuring long-term success. The second most described factor was continuity; especially in commitment of personnel and resources. The institutions spoke of advocacy in different ways. OhioLINK cited the value of gathering a large number of institutions to build a federated environment and success. NPS and AFIT both touched on the importance of aligning services with the strategic plan of the overarching institution. NPS also noted a “build-it and they’ll find it” approach to introducing systems which might not be directly on the users’ horizons but whose utility will become apparent upon discovery by the community, and also mentioned the need to ensure properly formatted metadata to prevent duplication. In gaining the support of the faculty to populate IR systems, WSU noted the value of extant librarian-faculty relations in creating a trusted environment for transition.

4.3.2.4 Role Evolution and Infrastructure

The advent of new technologies certainly affected the roles of the staff of these institutions, as well as their infrastructural requirements. Role-wise, the institutions all mentioned the need for computer literacy and familiarity. OhioLINK has spent significant effort in developing liaisons in its member institutions to enhance the exchange of information and system growth. Wright State University's librarians are versed in a full range of interactive software to work with their patrons using the methods comfortable to them, including chat, e-mail, instant messaging, and other rich interfaces. NPS cited workflow and the ability to communicate rapidly and effectively as necessary skills for its staff. MBARI noted a need for the librarian to be able to oversee IT personnel have a complete understanding of the systems being implemented. PAOI's librarian was self-instructed in the development of their entire catalog system, and grew it from the ground up from a few shelves to thousands of volumes. Other described roles were the ability to instruct faculty and students in the use of new systems, and the need to advocate for their use where and when appropriate. In all cases, modern librarian skill sets were described as requiring a high degree of information technology proficiency.

As they had with roles and responsibilities, all interviewed institutions remarked on the need to build up infrastructure to accommodate the introduction of new systems. Most described the buildup as smooth and gradual, involving expansion of servers and databases and the incorporation of new software packages as they were developed. WSU commented on increased bandwidth and electrical power needs, as well as the development of more dynamic web servers to support some of the new functionality being introduced.

4.3.2.5 Planned System Enhancements

When asked whether or not they planned enhancements to their IT systems, the universal response from all those interviewed was yes. Six of the eight institutions expressed a desire to extend new services to its users by adding more remote services, federated searching, improving interlibrary loan, and so forth. DLI is considering the addition of .mil wireless functionality to better serve its patrons. NPS is working on including mobile device applications to make their systems more accessible to remote users. WSU, NPS, and AFIT are all in the process of evaluating and developing their IR efforts to provide their users the content they seek, and OhioLINK is continuing to improve the Digital Resource Commons.

As far as concerns the customer base, the education and research/education focused institutions all expressed some idea of continuing to expand access to new users. The DoD institutions placed the additional caveat of remaining within mission, but were all willing to provide services to better service their distance learners (AFIT and NPS) as well as the DoD community at large. OhioLINK operates as a public good and focuses its energy on supporting not only its member institutions, but other libraries and institutions within the state as well. Other prevalent incremental improvements were continuous evaluation and redesign of library websites to account for changing patron behavior and ease of use, upgrading of library ILSs, and additional content.

4.3.3 Library Unplanned Activities

4.3.3.1 Other Technologies and Opportunities

In all cases, the institutions were aware of emerging technologies other than those already formally planned for that would improve the experiences of their respec-

tive customers. Most prevalent among these were those that would offer expanded access to their identified customers (mobile device tools, increased digital content, remote services, etc) and those that would offer richer media to their patrons, such as enhanced multimedia delivered through the website (for example a YouTube-like interface for video content). OhioLINK in particular made a distinction between institutional repositories and digital commons, noting that default installations of software like D-Space tend to be archival in nature and unconcerned with presentation; the commons perspective is to provide a rich interactive experience to knowledge seekers along with the archival function supported by repositories.

Also mentioned were more backend-type IT packages designed to help the libraries manage their collections and support of the various services they wished to offer. Some examples were content management services to help centralize and distribute digital content from across the institution, software to allow for seamless federated searching across multiple databases, and imaging software to allow for daily reimaging of terminals to reduce security concerns that otherwise prevent expansion of services.

These technologies were identified both proactively and reactively – proactively through multi-library meetings, conferences, vigilance on the part of IT personnel, and querying their customers; and reactively through responses to customer inquiries and identification of gaps in current service or inability to answer specific questions.

4.3.3.2 Barriers to Implementation and Success

The most popular answer to the barrier question was funding, irrespective of the categorization schema. Funding impact included both insufficient budget and the fact that the funding cycle often lags behind the introduction of new technology. Intellectual

property was mentioned by one institution inasmuch as it affects what sort of material can be shared on the web. For the DoD institutions, security was a major concern, both from the aspect that meeting requirements limits the range of potential services the institution can offer, and the challenge in developing systems to work with limited distribution and classified material to support customers operating in those areas.

WSU and OhioLINK both mentioned lack of contributions as barriers to the success in IR efforts, underscoring the need to continue to build advocacy and advertise the efficacy of the system. Personnel development was mentioned in regard to collaborative efforts between OhioLINK and member institutions. Should a member organization frequently replace liaison personnel, the time required to spin up a fresh liaison is time intensive, and in the mean time the quality of the product could suffer. NPS also mentioned the fact that there are hard skills necessary to establish and operate an institutional repository. They further described the necessity of developing good service-level agreements with the IT divisions to ensure continued support. Stability in roles and responsibilities goes a long way towards maintaining the quality of the product.

Some other concerns included lack of control over resources, both machine-oriented (i.e. control over servers) and personnel-related (control over IT professionals responsible for system development). Lack of personnel control can result in time delay and confusion when developing new systems, and lack of server control leaves one at the whim of the host, such that if the host server is incapable of handling certain applications or the host no longer wishes to support those applications, operations of the tenant are adversely affected. In concert with these items, geography itself was noted by WSU as a limiting factor, considering the increased amount of time needed for coordination

between dispersed collaborative units. Thus control and authority were described universally as affecting the agility of libraries to respond to and implement changes in technology that would benefit their patrons.

A final remark applicable to ownership deals with software support by vendors. NPS experienced the effects of a supplier's unwillingness to support legacy applications. When using proprietary software, there is always a risk that a publisher will no longer maintain support older systems which leads to their obsolescence.

4.3.3.3 Other Thoughts

The final question in the interview script encouraged open discussion of any other thoughts that might be relevant but which were not directly asked during the course of the session. The following are some noteworthy ideas that came out of this discussion. The Patriarch Athenagoras Orthodox Institute related some of the issues inherent in collecting abstruse texts. Due to its focus on Orthodox Christian manuscripts written well before any standard organizational system (such as the Dewey Decimal or Library of Congress numbering systems), and often coming from out-of-the-ordinary publication sources, categorizing and metadata population are more difficult, as is collection development itself, which requires old-fashioned scouring of dusty shelves in used bookstores. The Library of Congress system proved very flexible for accommodating these works into the collection. Both the PAOI and DLI noted the difficulties inherent in developing catalogs and retrieval systems for non-roman alphabets. These issues affect these institutions due to their large volume of holdings of foreign-language texts, especially those written in non-phonetic alphabets. UC Berkeley was also alluded to concerning its own trouble cataloging its extensive east-Asian collection, and this problem deepens when

attempting to expand access to a collection in large databases with differing alphabetizing structures. Partnering and collaboration were mentioned here as well by OhioLINK and all the research-focused institutions. This theme will be explored in more detail in the next section.

4.4 Themes and Trends by Category

4.4.1 Focus

From among the institutions sampled, institutional focus related to the types of services offered and holdings maintained. The educationally-focused institutions all held a wide range of media, including books, reference materials, multimedia, and various other mission-related texts. They also provided study space, areas for collaboration and relaxation, and various IT services such as internet access, e-mail, word-processing and other productivity applications, and so forth. External use revolved around the discovery of texts for circulation.

The research-focused institutions interviewed in this study tended to maintain large sets of subscription journals, theses, dissertations, and other archival-quality works produced by their host institutions. Enhancing source location was a prevalent theme in this area, with many libraries reflecting on their desire to provide knowledge seekers with the best information possible. All of the research institutions cited proxy or VPN authentication for the purpose of accessing online subscription journals and access to the catalog for the purpose of source location, consistent with the aforementioned desire to provide researchers with rapid access to the information they need to conduct their investigations.

4.4.2 Jurisdiction

Jurisdiction related to services offered as well. Security concerns in DoD institutions often obviate providing certain services such as commercial Wi-Fi which would otherwise benefit its customers. Another unique element of DoD affiliation was the desire to harness the power of some of these new technologies to develop systems to offer limited-distribution information in the correct venue. A theme for these institutions was the conflict between the .edu and .mil domain – the .edu provides certain collaborative advantages with other universities but makes some interoperability with .mil systems more challenging, while the .mil provides better functionality and security within the .mil but has certain limiting constraints in regard to software, management, and outreach and recognition among other universities.

4.4.3 Institution Size

Institutional size related to the ability of an organization to provide data and metrics for analysis. Larger institutions have more fragmentation of disciplines, making the gathering of certain statistical data much more of a challenge. Both Berkeley and WSU cited difficulty in finding any one person who could answer all of the questions posed in the questionnaire, and one potential institution was unable to participate citing an inability to gather the requisite information in time to respond for this study. The mid-sized institutions generally organized responses from one or two additional librarians; whereas the smaller institutions were generally able to readily respond to all areas of the questionnaire with little difficulty (in this case OhioLINK is grouped with small institutions due to the size of its staff).

Notwithstanding data collection issues, the large, public institutions (OhioLINK included) were in general responsible for the stewardship of a large body of the legacy of their universities and states. The universities house important works of literary and historical significance (such as Wright State's management of various media related to the Wright Brothers' accomplishments), and their respective collaborating repositories (CDL and OhioLINK) maintain large volumes of knowledge on behalf of the entire state and many member institutions. This is analogous to projects in Australia and the UK which often require contributions to the respective national repository for funding support, or DTIC to which completed DoD project reports must be submitted at completion or contract close-out; only without the government mandate for submission which has left these IRs functioning on a more voluntary basis.

4.4.4 All Categories

Funding and budgets were mentioned as concerns across-the-board; more generally, resources, including personnel and control over how work is accomplished by various divisions were noted as important to pursuing these initiatives. Institutional repositories are expensive to build and operate. According to the Institutional Repository SPEC Kit, a survey conducted in 2006 among 123 members of the association of research libraries, the mean cost of implementing an IR was \$182,550, and the mean annual operation cost was \$113,543 (Bailey, Institutional Repositories SPEC Kit, 2007). The major facets in the personnel issues revolve around *time*: the time it takes to train proficient liaisons to interact with complex repository systems, the time it takes to resolve labor issues associated with the roles and responsibilities of various personnel with valued skill sets – especially contractor personnel who are not explicitly given contractual au-

thority to perform certain roles for which they are otherwise capable, and the organizational structure of a library's host institution where it impacts the library's ability to rapidly and successfully respond to the advent of new technologies and implement systems. The results of the IR SPEC Kit Study support this as well, stating that the majority of implementers of IR technology had an IT function within their direct reporting chain (Bailey, Institutional Repositories SPEC Kit, 2007). Given the rapid evolution and maturation of information technology, timely incorporation of valuable systems is vital to a library's ability to stay on the leading edge of service to its customers.

Time and budget also play a role with respect to infrastructure. Budget often affects where servers and IT personnel are located, and those libraries without direct oversight of their web or data servers or computer support all related some level of concern over their ability to affect the quality and modernity of that infrastructure. In one case, a library's new ILS was too powerful for the server it was forced to reside upon, producing http errors for certain functions and a general reduction in the quality of service that system was intended to provide.

Open metadata architecture was another universal theme, with all six of the respondents familiar with IRs gravitating to this standard. The institutions all conveyed the importance of ensuring proper population of metadata to improve the outreach of the reposing knowledge, with several dedicating (or planning to dedicate) metadata librarians to ensure this was the case. Adherence to standards is generally enforceable through the submissions process, as is quality control over content for the partnered institutions.

In regard to collaboration and partnering, all of the institutions interviewed mentioned some form of initiative to extend the reach of their services or collections. All of the libraries mentioned some level of interlibrary loan as a service they either provided or planned to provide, and most listed or planned to list their collections with OCLC to enhance global outward access. The University of California is currently in the midst of merging its venerable Melvyl catalog into the OCLC's WorldCat, providing more seamless access to its works from the at-large community.

The existence of OhioLINK and the other regional and subject repositories described throughout this study and the sizes of their service populations (OhioLINK supports 90+ member institutions and other, non-member organizations as well) are all evidence of a trend towards the pooling of resources and centralizing of information, however, the preponderance of those interviewed maintained their own systems for exchanging certain forms of content for various reasons.



5 CONCLUSIONS AND LIMITATIONS

5.1 Conclusions

5.1.1 What are Libraries Doing Today?

This is a time of transition for many institutions, and that transition appears to be following the overall trend of the world to become more interconnected. Institutional repository technology is indeed on the lips of a growing number of academic and research libraries, with all of the research-oriented institutions in this study in some stage of investigating or implementing some aspect of the technology.

There is certainly a trend towards partnering and collaboration. Centralized repositories provide research institutions a way to archive their intellectual produce without the need to build up costly infrastructure to support them; however, such systems may impose certain limitations on member institutions. Centralizing resources creates a dependence on the hosting authority and may limit the type of material the client is able to load. A system like DTIC, for example, with very clear criteria for the sorts of reports it allows, may not meet the needs of an institution seeking to publish faculty CV links or to share and steward working papers and other forms of grey literature. In response to these limitations most of the institutions interviewed defined a more hybrid approach, where the material acceptable to the repository was loaded therein, and that material which was not is provided by some other means, either through separate repository systems or via the institutions' websites.

5.1.2 What have Libraries Planned?

Academic libraries are, as ever, service oriented, and are always examining ways to enhance the products they offer their clientele. They are all immersed in the technologies that allow them provide the best information for their patrons and to that end all of the institutions represented in this study had various continuous improvements and enhancements planned over the next few years.

Even with the advent of institutional repositories, journals don't seem to be going away for the present. Proxy access to subscription journals was a theme across the board for research organizations, and if anything making this access easier to obtain through upgrades was the only direction on that topic. Various forms of grey literature were discussed by a number of institutions, all of which planned their own IR or an ancillary system in parallel to a regional or subject repository. Further convoluting the above is the fact that there even within institutions championing open access there remains some conflict between open journals and subscription issues. While, for example, the University of California's office of scholarly communication is very clear on its stance in favor of open access, member universities maintain journals such as the California Management Review that are distributed via paid subscription only (The Regents of the University of California, 2009).

A major theme throughout the entire interview period was one of seamless interfaces. Those institutions that are planning for or in the process of implementing repositories through partnerships have all expressed a desire to do so in a manner that is transparent to their customers. The regional repository systems spoken with or about tend to support these interfaces – DTIC is working with NPS to develop a feed to enhance direct

sharing and metadata coordination, and AFIT plans its own seamless DTIC interface to provide its customers rapid access to their vast collection. According to its strategic plan, DTIC is itself interested in providing easy access to its stewarded works to that end is examining ways to simplify user access through LDAP authentication, CAC login, and other single-sign-on technologies (The Defense Technical Information Center, 2008). In providing its member institutions with their repositories, OhioLINK makes a concerted effort to build an interface that uses the livery and style of those institutions' websites. Similarly, the integration of Melvyl onto WorldCat retains Melvyl's standard iconography and field-controlled access to the UC collections, but opens access to the world's collections as well. All of these examples demonstrate a commitment on the part of these libraries to provide their patrons rapid access to the knowledge they need in as simple a manner as possible.

The issue of livery is interesting from the perspective of the roles of institutional repositories and the conflict between centralized and institution-based systems. The reader will recall that IRs fulfill a role with respect to institutional identity – that one of the reasons their proponents cite for using them instead of journals is that journals enhance their own prestige while diluting the contributions of any particular institution, and thus the IR provides an institutional record of its own intellectual accomplishments. Centralizing a repository might also appear on the surface to dilute the accomplishments of its contributors; however, the addition of branding promotes the member institution and presents a medium between locally- and remotely-hosted repositories – a vision OhioLINK refers to as “your files, our server, your site.” Consistent with this theme, the aquatic commons allows for browsing by subject *and* by agency, so that a knowledge

seeker interested in the collective impact of a given center on the repository can easily come by that knowledge as well as the total number of contributions made to the repository by each.

5.1.3 What Unplanned Activities do Libraries Find Important?

The provision of rich content and experience was another universal desire of all of the centers sampled in this study. Simply allowing multimedia files to repose on a database was often deemed insufficient; rather presenting those files in a pleasing, interactive way to their patrons was the more preferred approach, and continuing to evaluate and harness the technologies that enable that to happen was something of a priority in many cases. You-Tube like interfaces, image manipulation, and integrated audio were all provided as examples of interaction tools that modern, tech-savvy customers would appreciate both during data gathering and leisure use.

Mobile device tools, chatting and instant messaging sessions between patrons and library staff, and various other modes of communication were all mentioned as services that are geared to provide rapid, enhanced access to all forms of media that a user could want, and in a manner as comfortable to that user as possible. These technologies address individual user dispositions and needs by providing variety in interfacing with the library to conduct their business.

Enhancing backend software to augment the fulfillment of the above services and other as-yet undiscovered tools was a prevalent desire as well. All libraries described a proactive approach to assessing the needs of their patrons through surveying and colloquia and to continuously assessing the state of the practice through inter-library and society conferences, regular peer interactions, and so forth. All of this further underscores

that libraries are active at the forefront of information technology and mindful of new ways to incorporate emerging tools into their operations.

The overall impression conveyed by the results of this study is that academic libraries are in the midst of transition towards a richer, more collaborative operating environment. Institutional repositories, federated searching, and various other web utilities are methods being explored and implemented as these libraries continue their drive to increase the reach and quality of their services, and their ability to safeguard the intellectual produce of their host institutions. Libraries are cognizant of the needs and desires of their using community, and continue to find new ways of utilizing available information technology to provide that community the best products they can. Libraries remain on the cutting edge of knowledge production and management as they have been since the times of Ptolemy, and should continue to be so as long as they maintain their technological proactivity and expertise in information science. Though the term “institutional repository” does have some elements of a catch-phrase, the responsible implementation of the aspects of these systems that support the faculty and students of the sponsoring organization does present an opportunity to expand the reach of an institution’s works and the faculty’s ability to communicate. Partnering offers certain opportunities but may come at the expense of some content. In striking the proper balance it is very important that a library interested in pursuing a repository conduct some form of audit of its faculty’s needs and determine the best way to meet those. Hybrids between centralized and localized repositories provide one avenue for resolving this conflict – a needs assessment will go a long way towards determining what the faculty needs are and how best to meet them.

That said, however, a library should not allow itself to become over-constrained by such an assessment. In keeping with the systems engineering principle of emergent behavior, there are times when it is impossible to anticipate the full range of uses a potential technology will provide, especially by those who have no experience with a new technological system. Until users begin to tinker with a system and to truly explore its capabilities, they may be unaware of uses that could completely alter the way they conduct their business in much the same way as e-mail altered the way businesses communicate or iPods the way people spend their down time. Ensuring the needs of the organization are met is an important activity, but failing to allow for growth because a needs assessment does not reveal an unknown need may cause the system to be redundant and less-well-used than it could be. Thus, in keeping with the historical mandate of always remaining at the forefront of knowledge and technology production, it is imperative that the library continue to scour the community for those technologies that provide the best potential to improve on and expand the services they offer to whatever extent allowed by their resources and individual missions.

5.2 Limitations and Considerations for Future Research

This study examined ongoing efforts by a small sample of libraries of various sizes, with the goal of providing the AFIT/AFRL library a thin cross-section of the experiences of libraries with similar and differing needs and situations to help guide their own IR development. To gain a deeper understanding of the state of institutional repository and other IT application and success, further study of a large sample of academic libraries in any of the categories this study targeted would be appropriate. The expan-

sion of this research could be accomplished in several ways. Targeting a specific category of interest from the above classifications or an entirely new scheme would give results that could be more tailored to a similar institution. Canvassing a large population with a similar distribution of categories would help to corroborate the findings of this study and help form a more general picture of the state of the practice. Another method of expanding this study would be to target regional and subject repositories themselves, due to their interaction with a large sample of institutions. This would narrow the scope to those institutions with some direct institutional repository experience, but would highlight the maturity of their efforts and barriers to implementation and use.

Snowball sampling, while effective at identifying institutions that the researcher mightn't have otherwise thought of, does have a potential drawback in that asking advanced or otherwise high-caliber institutions for suggestions on whom to visit is likely to yield other high-caliber institutions (certainly the level of excitement with which one is directed to another institution indicates the esteem felt by the referrer to the referee). If one desires a mix of more- *and* less-advanced institutions some different tactic might be in order, such as asking for a range of potential subjects based on the institution's understanding of their levels of maturity, or to start with a truly random sample drawn from across the nation.

Another facet of this issue that was out of the scope of this investigation but which might provide more insight into the long-term viability of IRs and other archival systems would be an assessment of how well repositories are accepted by different universities and research centers for the purposes of faculty standing and promotion. Since journal publication currently plays a very large role in how rapidly professors rise

through the ranks, determining whether or not works residing on an institutional repository carry the same weight as those in journals (among various classes of institutions) would provide a valuable insight into whether the four originating purposes of the technology are being fulfilled or whether these systems are serving a more archival function.

Those who might wish to expand this study through interviewing with larger libraries should be aware that, as with many large organizations, libraries of major universities and research centers are often broken down by discipline and focus. Additionally, these institutions may also have many libraries dedicated to various subjects, each with their own needs and challenges. When the structure of a library is thus fragmented, it becomes difficult to find a single individual capable of answering IT infrastructure, institutional vision, repository design, and library statistics all at once. This can lengthen the time of the response significantly as answers are sought between departments and may also reduce the ability of a participant to provide a complete answer. Some of these obstacles occurred during the course of this study; fortunately however the interview processes with the larger institutions were initiated early enough to result in data prior to the termination of the effort, though more time would certainly have produced more detailed resolution on some of the deeper questions.

Finally, I would recommend that anyone seeking to implement an institutional repository peruse Susan Gibbons' *Establishing an Institutional Repository*, as it covers many of the steps a library ought to go through before creating its own repository. Also, Charles W. Bailey's *Institutional Repositories SPEC Kit* gives a very good summary of the costs associated with developing and maintaining IRs as well as their implementers' opi-

nions on their level of maturity, data which could help a library decide on whether an IR would be a cost-effective solution for its institution's needs.

APPENDIX I: LIBRARY INTERVIEW QUESTIONNAIRE

1. Library Classification Questions

- 1.1. What is the approximate size of the population this library serves?
- 1.2. What percentage of that population is local, what percentage is geographically-separated (i.e. distance learners)?
- 1.3. What is the most popular method of locating media for local customers?
 - (a) Catalog
 - (b) Research Librarian
 - (c) Browsing the stacks
 - (d) Other
- 1.4. What is the most popular method of locating media for remote customers?
 - (a) Web-based catalog/retrieval system
 - (b) Other automated systems
 - (c) Human-operated telephone system
 - (c) Other
- 1.5. What is the most popular use of the library for local customers?
 - (a) Internet Access
 - (b) Source Location/Retrieval
 - (c) Study/Relaxation
 - (d) Other (Specify)
- 1.6. Describe the library's holdings: approximately how much of the library's material is
 - (a) Produced by the Host institution
 - (b) Subscription Journals
 - (c) Reference Material
 - (d) Books related to the institution's stated mission
 - (e) Other
- 1.7. What is the most popular form of media checked out to local customers?
 - (a) Books

- (b) Periodicals
 - (c) Audio Media
 - (d) Video Media
 - (e) Other (Specify)
- 1.8. What is the most popular use of the library for remote users?
- (a) Locally produced material
 - (b) Access to Subscription Journals
 - (c) Remote access to/checkout of library holdings
 - (d) Library-specific services (Specify)
 - (e) Other (Specify)
2. *What are libraries doing today to address and incorporate new technologies in serving their customers?*
- 2.1. Does the library use IT in its daily operations?
- 2.2. How long has the library used IT in its daily operations?
- 2.3. What form of IT services does the library employ, both internally (inventory management, cataloging, etc) and externally (web-based services for at-large community)?
- 2.4. What are the most popular IT applications for local users?
- 2.5. What are the most popular IT applications for remote users?
- 2.6. Is the library's collection linked to the collection of a parent or sister organization?
- 2.7. What systems do librarians use to locate holdings?
- 2.8. What systems do circulation librarians use to loan materials and manage inventory?
- 2.9. Does the library's host institution produce any amount of intellectual work?
- 2.9.1. Describe these works
- (a) Theses/ Dissertations
 - (b) Faculty CV Citations/Links

- (c) Faculty Journal Articles
- (d) Other (Specify)

2.9.2. Is the library responsible for stewarding any of these works?

2.9.3. What types of locally-produced works does the library hold?

- (a) Theses/Dissertations
- (b) Faculty CV Citations/Links
- (c) Faculty Journal Articles
- (d) Other (Specify)

2.9.4. How does the library make these works available to its customers?

- (a) Local Hardcopy
- (b) Journal Subscriptions
- (c) Locally-managed IT System
- (d) Linking to other IT Systems
- (e) Other

2.10. Does the library's host institution maintain an institutional repository or similar system?

2.10.1. Is the library or another division of the host institution responsible for maintaining this system?

2.10.2. Was this system home-grown or tailor-made?

2.10.3. Does the system use an open-source infrastructure (like D-Space) or a proprietary structure (like eScholarship)?

2.10.4. What manners of work reside on the system?

2.10.5. How are roles and responsibilities distributed?

2.10.5.1. Who contributes?

2.10.5.2. Who consumes?

2.10.5.3. Who manages the system?

2.10.5.4. Is the system moderated? By whom?

2.10.6. How do users learn to interact with the system?

2.10.7. How do contributors load content?

- (a) Directly?
- (b) Through an intermediary?

- 2.10.8. Does the system interact with external systems?
 - 2.10.8.1. Is this interaction through standards (such as OAI-PMH)?
 - 2.10.8.2. Is the knowledge in the system accessible to all potential consumers or does the system maintain an approved user list?
 - 2.10.9. Does the system enforce metadata standards?
 - 2.10.9.1. How are these standards enforced?
 - 2.10.10. How is metadata audited?
 - 2.10.11. How well is the system used?
 - 2.10.11.1. How is this use defined?
 - 2.10.12. Would you define the system as a success?
 - 2.10.12.1. What criteria do you use to define success?
 - 2.10.12.2. If the system is successful, in what ways is it successful?
 - 2.10.12.2.1. To what elements do you attribute this success?
 - 2.10.12.3. If the system is unsuccessful, in what ways is it unsuccessful?
 - 2.10.12.3.1. To what elements do you attribute this lack of success?
 - 2.10.12.4. Are there plans in place to improve the use of the system?
 - 2.11. Has the incorporation of IT systems into the library affected the roles of the staff librarians?
 - 2.11.1. In what ways have these roles changed?
 - 2.11.2. What education/training requirements were necessary to help the staff grow into these roles?
 - 2.11.3. Did this transition occur smoothly or were there major roadblocks?
 - 2.12. Was there any additional infrastructure necessary to institute the system?
 - 2.12.1. Did this transition occur smoothly or were there major roadblocks?
 - 2.13. Describe any other barriers to the successful implementation of the system and how they were overcome.
3. *What IT initiatives have libraries planned already?*
- 3.1. Does the library feel a need to enhance the services it offers its customers?
 - 3.1.1. How did the library come to recognize this need?

- 3.2. Does the library wish to enhance current services or extend new services to its customers?
 - 3.2.1. What types of services does the library plan to provide its customers?
 - 3.2.2. What types of content does the library plan to provide its customers?
 - 3.2.3. Has the library performed an audit to determine its culture (and that of the host institution), information holdings, and needs?
- 3.3. Does the library wish to maintain its current customer base or extend its services to *new* customers?
 - 3.3.1. To whom would the library like to extend its services?
- 3.4. Does the library plan to develop an IT system to better serve its customers (current or potential)?
 - 3.4.1. What manner of IT system does the library feel would best benefit its customers?
 - (a) Catalog/Indexing service
 - (b) Search Engines
 - (c) Knowledge-management-type system such as institutional repositories
 - (d) Other Services (Specify)
 - 3.4.2. Describe the system.
 - 3.4.3. For IR-Type Systems:
 - 3.4.3.1. Will the system be home-grown or tailor-made?
 - 3.4.3.2. Will the system use an open-source infrastructure (like D-Space) or a proprietary structure (like eScholarship)?
 - 3.4.3.3. What manners of work will reside on the system?
 - 3.4.3.4. How will the roles and responsibilities be distributed?
 - 3.4.3.4.1. Who will contribute?
 - 3.4.3.4.2. Who will consume?
 - 3.4.3.4.3. Who will manage the system?
 - 3.4.3.4.4. Will the system be moderated? By whom?
 - 3.4.3.5. How will users learn to interact with the system?

- 3.4.3.6. How will contributors load content?
 - (a) Directly?
 - (b) Through an intermediary?
- 3.4.3.7. Will the system interact with external systems?
 - 3.4.3.7.1. Will this interaction be through standards (such as OAI-PMH)?
 - 3.4.3.7.2. Will the knowledge in the system be accessible to all potential consumers or will the system maintain an approved user list?
- 3.4.3.8. Will the system enforce metadata standards?
 - 3.4.3.8.1. How will these standards be enforced?
- 3.4.3.9. How will metadata be audited?
- 3.4.3.10. Has the library developed success criteria?
 - 3.4.3.10.1. Has the library identified potential sources of failure?
 - 3.4.3.10.2. How will the library act to ensure success?
- 3.4.4. Does the library foresee the roles of the staff librarians being significantly altered with the introduction of the new system?
 - 3.4.4.1. How will these roles change?
 - 3.4.4.2. What education/training requirements does the library feel will be necessary to help the staff grow into these new roles?
 - 3.4.4.2.1. What roadblocks does the library envisage in accomplishing this development?
 - 3.4.4.3. Will there be any additional infrastructure necessary to institute the system?
 - 3.4.4.3.1. What roadblocks does the library envisage in accomplishing this development?
 - 3.4.4.4. Describe any other identified potential barriers to the successful implementation of the system and how the library plans to overcome them.

4. *Beyond what you're doing or what you're planning to do, what do you see on the horizon that may be important to better serve their customers and make the information the library stewards available?*
- 4.1. Is the library aware of other technologies that could enhance its ability to serve its institution and customers?
 - 4.2. What types of technologies does the library feel would offer the most potential in serving the customers?
 - 4.3. For what reasons?
 - 4.4. What barriers does the library feel exist that might hamper the successful implementation of such a new system?
 - 4.5. In addition to questions I've already asked, is there anything else that I should know about this topic that would help me better understand the way that library services are going in the future?

APPENDIX II: SUMMARY RESPONSE TABLES

Library Classification

Question	OhioLINK	UC Berkeley	WSU
1.1 Library Customer Population	600000	37437	16000
1.2 Distribution (Local/Distance)	100% D	No Stats	No Stats
1.3 Local Media Location	N/A	Pathfinder/ MELVYL	Online Cat/ Website
1.4 Remote Media Location	Google/Web Ifc	Website	Website
1.5 Local Library Use	N/A	Internet/ Source Location/ Study & Relaxa- tion/ Collaboration/ WiFi/ Applications/	Internet/ Source Location/ Study & Relaxa- tion/ Collaboration/ WiFi/ Applications
1.6 Library Holdings	Theses/ Dissertations/ Archival	Books/ Journals/ Theses/ Dissertations/ Archive	Books/ Journals/ Theses/ Dissertations/ Archive
1.7 Popularly Circulated Media	Books (Via ILL)	Books	
1.8 Remote Library Use	Journals/ Catalog/ Repository	Proxy access to Subscription Jour- nals	Proxy access to Subscription Jour- nals

Table A-1: Summary Responses to Classification Questions I

Question	DLI	NPS	PAOI
1.1 Library Customer Population	4500	2500	1450
1.2 Distribution (Local/Distance)	100% L	75% L	95%L
1.3 Local Media Location	Online Cat	Online Cat/ Website	Online Cat/ Browsing
1.4 Remote Media Location	Website/ILL	Website	GTU Website
1.5 Local Library Use	Internet/ Source Location/ Study & Relaxa- tion/ WiFi	Internet/ Source Location/ Study & Relaxa- tion/ Collaboration	Source Location/ Study & Relaxa- tion/ Collaboration
1.6 Library Holdings	Books/ Course Materials	Books/ Journals/ Theses/ Dissertations/ Archive	Books/ Journals/ Lecture Series
1.7 Popularly Circulated Media	Books/ Multimedia	Books/ Some Multimedia	Books/ Journals
1.8 Remote Library Use	Online Databases	Proxy access to Subscription Jour- nals/ Online Databases	Book Location

Table A-2: Summary Responses to Classification Questions II

	Question	AFIT	MBARI
1.1	Library Customer Population	750	500
1.2	Distribution (Local/Distance)	90%L	100%L
1.3	Local Media Location	Online Cat	Web Ifc for Lib Holdings Video Ifc for Media
1.4	Remote Media Location	Website/WebFeat	Website for media
1.5	Local Library Use	Internet/ Study & Relaxation/ Collaboration/ Applications	Source Location
1.6	Library Holdings	Books/ Journals/ Theses/ Dissertations	Books & Tech Reports/ Video/ Journals
1.7	Popularly Circulated Media	Books/ Some Multimedia	Books
1.8	Remote Library Use	Proxy access to Subscription Journals	Proxy access to Subscription Journals

Table A-3: Summary Responses to Classification Questions III

	Question	OhioLINK	UC Berkeley	WSU
2.1	Does the Library Use IT	Yes	Yes	Yes
2.2	For How Long	1992	1980	1987
2.3	Form of IT Services	Digital Resource Commons/ Repository/ Federated Search	ILS (GLADIS/Pathfinder - Millenium)/ Catalog/ Circulation/ Acquisitions/ Digital Object Content Management/ Support Applications	ILS (III)/ User Applications/ Course reserves/ Epoxy (Proxy)/ Purchase software
2.4	Popular Local IT Apps		No Metric	Catalog/Local Re- sources ILS/Browsers
2.5	Popular Remote IT Apps		No Metric, No Distinc- tion from Local	Encore/Wright Cata- log Browsers/EZProxy
2.6	External Linkage		CDL, Other UCs, Stan- ford	OhioLINK, OCLC
2.7	Holding Location		GLADIS, Pathfinder (UCB) MELVYL (UC) Transitioning to Worldcat	ILS University Catalog
2.8	Checkout Software		GLADIS - Homegrown	ILS Docutek for ereserves ILLiad for inter-lib loans
2.9	Intellectual Work Production	No - stewardship of: Theses/Dissertations/ Archival media of member institutions	Yes Thes- es/Dissertations/Article s & Links	Yes Thes- es/Dissertations/Archiv es
2.10	∃ an IR or similar system	Yes	Yes - CDL Provides these services	Yes - OhioLINK & Local
2.11	Has incorporation affected roles	Yes	Yes	Yes - Librarians use rich media interfaces to exchange with patrons
2.12	Was infrastructure necessary?	Yes	Yes - Gradual Build-up of Infrastructure	Yes - more dynamic web hosting to support pooling, etc More bandwidth and electric power
2.13	Describe Barriers to implementation	Populating Data/ Funding/ Developing Personnel	Underpopulation - advocacy requirement	Funding Reliance on external hosting

Table A-4: Responses to Library IT Initiative Questions I

	Question	DLI	NPS	PAOI
2.1	Does the Library Use IT	Yes	Yes	Yes
2.2	For How Long	1990s	1980s	1990s
2.3	Form of IT Services	ILS (TLC)/ User Applications	ILS (SIRSI)/ Online Catalog, Circulation, Discovery, Inventory Management/ Print Servers/ ILL service/ Proxy	ILS (Filemaker)
2.4	Popular Local IT Apps	Service Websites Commercial E-mail	No Metric	Filemaker
2.5	Popular Remote IT Apps	AKO Material and Web Catalog	Resource Pages, Teaching Calendars, Library Locator	
2.6	External Linkage	OCLC	OCLC, Collection is stand-alone	GTU, OCLC
2.7	Holding Location	OCLC or individual catalog ILS	ILS (SIRSI)	Filemaker Catalog
2.8	Checkout Software	ILS	ILS Includes open and restricted instances	Filemaker Catalog
2.9	Intellectual Work Production	Course Development/ Curricula	Yes Theses/Dissertations/Tech Reports/Publications	Yes Lecture Series
2.10	∃ an IR or similar system	No - Institutional Archive	Yes - Called Digital Rep	No
2.11	Has incorporation affected roles	Yes	Yes - workflow, discovery, and access - greater communication	Yes - Librarian was self-taught in system development
2.12	Was infrastructure necessary?	Yes - data server hosted by DLI, webserver hosted by library	Yes - smooth buildup of new systems	Yes - 1 workstation and network interface
2.13	Describe Barriers to implementation	Lack of data server ownership is a limiting constraint	Lack of publisher support	Funding

Table A-5: Responses to Library IT Initiative Questions II

	Question	AFIT	MBARI
2.1	Does the Library Use IT	Yes	Yes
2.2	For How Long	1980s (DTIC in the 70s)	1995
2.3	Form of IT Services	ILS (Innovative Interfaces) MS Office	ILS (EOS)/ 3-part GUI for video files Multiple content areas on website Shibboleth (Proxy)
2.4	Popular Local IT Apps	MS Office Suite	Zotero
2.5	Popular Remote IT Apps	VPN for Proxy Access	VPN at MBARI Shibboleth (Proxy)
2.6	External Linkage	AFRL is Co-Located	OCLC
2.7	Holding Location	ILS (First SIRSI, now Innovative Interfaces)	EOS Browsing
2.8	Checkout Software	ILS	EOS
2.9	Intellectual Work Production	Yes Theses/Dissertations Faculty Journal Pubs Faculty-Published Books Tangibles	Yes Theses/Dissertations from Moss Landing Side MBARI Patents/Inventions/Illustrations
2.10	∃ an IR or similar system	Not yet - Working on linking to DTIC	Yes - Repository Through IAMFLIC Local querriable database (over 1M files)
2.11	Has incorporation affected roles	Yes - Computer proficiency	Yes - increased responsibilities, expertise, supervision of IT folks
2.12	Was infrastructure necessary?	Yes	Yes
2.13	Describe Barriers to implementation	Funding Separation of powers	Funding

Table A-6: Responses to Library IT Initiative Questions III

	Question	OhioLINK	UC Berkeley	WSU
3.1	Does the Library Plan Enhancements	Yes	Yes	Yes - website redesign to enhance ease of use
3.2	Enhance/Extend	Enhance	Both	Both
3.3	Maintain/Extend Customers	Both	Both	Both - drive to include more off-campus users
3.4	Further Develop IT	Continuous improvements to DRC	Millenium (ILS)/ Digital Content Management	IR/ New Website

Table A-7: Responses to Library Planning Questions I

	Question	DLI	NPS	PAOI
3.1	Does the Library Plan Enhancements	Yes - as they come along	Yes - a new ILS	Indirectly - library is migrating to GTU
3.2	Enhance/Extend	Both - add .mil for wireless	Both - mobile tools, federated searching, and the IR	N/A
3.3	Maintain/Extend Customers	Remain within mission, but ILLs are fine; collaboration with Stanford	Primarily serve NPS, provide better service to distance learners Extend to DoD customers, distributed environment	Both - Increase worldcat functionality
3.4	Further Develop IT	Nothing specific - lots of possibilities including full text	IR	N/A

Table A-8: Responses to Library Planning Questions II

	Question	AFIT	MBARI
3.1	Does the Library Plan Enhancements	Yes - desire to always be able to help patrons	Yes
3.2	Enhance/Extend	Both - extend to distance learners, get back into ILLs, set up easy proxy for offsite users, retrospectives, and IR	Enhance - improve Shibboleth VPN for security/ease of use
3.3	Maintain/Extend Customers	Both - extend to distance learners and alumni	Maintain
3.4	Further Develop IT	Yes - create homegrown interface to DTIC	No major new IT development

Table A-9: Responses to Library Planning Questions III

	Question	OhioLINK	UC Berkeley	WSU
4.1	Is the library aware of other applications?	Yes View of the DRC as a commons, rather than simply a repository	Yes Enhanced collection management to centralize campus-wide digital content Backend support software	Yes IT folks are ahead of the curve, always examining new things
4.2	What has the greatest potential	Richer media, streams (YouTube like experience), image manipulation, etc	eBooks, streaming media, etc	Adding audio and other content
4.3	Reasons	D-Space is archival, does not care about presentation, commons are concerned with presentation and experience	Users want convenient access to content	Customers have expressed desire for that type of content
4.4	Barriers to Implementation	Resources - small staff, time to spin-up new liaisons, funding	Copyright	Budget Withdrawal of support by OhioLINK
4.5	Other Thoughts			

Table A-10: Responses to Potential Applications I

	Question	DLI	NPS	PAOI
4.1	Is the library aware of other applications?	Yes More online content	Yes Mobile technologies, content management, classified, and FOUO content	Yes
4.2	What has the greatest potential	Remote customer services Regaining full text	Mobile and Federated Searching	Full merge with GTU library
4.3	Reasons	Internet librarian meetings identify new tools to enhance services	Info security/assurance	Better protection of texts, accessibility through GTU web infrastructure
4.4	Barriers to Implementation	Budget Security issues with .mil domain	Budget POMs're slow People with hard skills	Budget Conflict between retaining text and centralizing at GTU
4.5	Other Thoughts	Issues with non-Roman alphabets	Assessment/return on investment Marketing Partnerships are big	Issues with non-Roman alphabets Collecting abstruse works

Table A-II: Responses to Library Planning Questions II

	Question	AFIT	MBARI
4.1	Is the library aware of other applications?	Yes Commercial WiFi Opening Terms to commercial e-mail	Yes The role IAMFLIC plays to coordinate marine laboratories
4.2	What has the greatest potential	Deep-freeze software to reimage machines each night	
4.3	Reasons	Mitigates security concerns to allow AFIT to provide "normal" library services	Identified through local library meetings and marine lab meetings
4.4	Barriers to Implementation	Access to programmers in SC & AFRL side of lib Security issues Staffing & Continuity Resistance to change	
4.5	Other Thoughts		Partnering with IAMFLIC

Table A-12: Responses to Library Planning Questions III

	Theme Area	OhioLINK	UC Berkeley	WSU
IR.1	Recognizing the need	Created to share Physical resources	Attention to Faculty and student comments Web Comments Interviews with Faculty Focus Groups Conferences Awareness of other Libraries' activities	Focus groups Online surveys pushed to students and faculty Conferences
IR.2	Configured off-the-shelf systems	Stand-alone D-Space Virtual Machines	eScholarship CDL, no additional IR	OhioLINK - configured D-Space environment
IR.3	Available content	Theses Dissertations Archival Works	OAC eScholarship Preservation Repository	EDT: Theses & Dissertations DRC: Intellectual Legacy
IR.4	Roles and responsibilities	Contributors: University Liaisons Consumers: All Mgmt: Local VMs No Moderators except bulk submission Content loaded by Liaisons Access control by IP/proxy	Reside with CDL Collection management resides with library as backend software Access control by IP/proxy	Contributors: Faculty Consumers: All Mgmt: WSU manages content, OhioLINK the system Moderated through staff, eventually LDAP for submission and lib staff review Content loaded by staff Access control by IP/proxy - allows for secure apps
IR.5	Quality control, metadata, and interaction	OAI-PMH PKP Harvester Metadata audited by liaisons		Dublin Core for Metadata, with OhioLINK mods Enforcement by staff at entry - metadata librarian & staff review
IR.6	Success and success criteria	Use defined by hits: Google Analytics Success by capability: does system deliver promised capability Currently Incomplete	Customers receive rapid access to digital content	Use defined by hits/click-throughs: Google Analytics Success also by content population Early to judge success - fresh system
IR.7	Ensuring Success	Well-trained Liaisons No national requirement makes it harder - gathering institutions together builds federation and success		Advocacy Building among faculty by librarians Adding features not necessarily offered by D-Space

Table A-13: Responses to IR-Related Questions I

	Theme Area	DLI	NPS	PAOI
IR.1	Recognizing the need	Observed reduction in collaboration due to .mil restrictions Restrictions also affect the ability of ILS software to function Conferences Annual tech audits to see what exists	Focus groups in the late 90s New assessment tools forthcoming	Direct interaction with patrons
IR.2	Configured off-the-shelf systems	Vendor-supplied ILS - no IR planned or extant	Currently Proprietary - Hyperion Moving to Equella or D-Space	Homegrown System for cataloging
IR.3	Available content	N/A	Old: Theses & Dissertations, NPS Pubs, Dept Reports, Joint-pubs, Congressional research docs New: Add Grey lit, ppts., databases & journals, NPS scholarly content	N/A
IR.4	Roles and responsibilities	N/A	Old: Contributors: Students, Faculty, Departments provide content Consumers: All for unclass system Mgmt: Library manages content, sys manages infrastructure Content loaded by staff New: Workflow, content loaded through blackboard Access control by IP/grouping	N/A
IR.5	Quality control, metadata, and interaction	N/A	Old: No OAI-PMH compliance, Exposure through Google Apps Enforcement through templates and mediation Audit through OCLC reports New: OAI-PMH, Workflows, submission reviewed prior to final post Library ensures QC	N/A
IR.6	Success and success criteria	N/A	Use defined by hits Success defined by ease of use for staff, availability & reliability for users	N/A
IR.7	Ensuring Success	N/A	Good Metadata, lack of duplication Dedicate staff to managing content - can't rely on community Provide carrots - repurposing content good Advocacy Aligning services with the strategic plan	N/A

Table A-14: Responses to IR-Related Questions II

	Theme Area	AFIT	MBARI
IR.1	Recognizing the need	Desire to answer patrons' questions Graduate Survey, Climate Survey, Main Survey Faculty needs assessment	All-hands brown bag lunches Informal gatherings among bay-area librarians
IR.2	Configured off-the-shelf systems	Interface with DTIC	EPrints 3 (according to website and interview)
IR.3	Available content	Theses, Dissertations, Tech Reports, faculty CV, graduate research projects	Dive data - images, video, ancillary data (lat/long/etc), mapping, etc off of MBARI website IR hosts members' scholarly materials, published and unpublished
IR.4	Roles and responsibilities	Contributors: Faculty and Students Consumers: Faculty, Students, Outside Researchers, open to all via DTIC content management Mgmt: Library manages interface Access control by IP/proxy, DTIC access control protocols	Contributions to repository through IAMSLIC (multinational) Contributors: Members at will Consumers: All (through google) Mgmt: Florida Center for Lib Automation Moderated by IAMSLIC editorial board
IR.5	Quality control, metadata, and interaction	Enforcement through submission, DTIC processes Interaction through seamless interface, DTIC, or the web	IAMSLIC board does QC Interaction through Google or Aquatic Commons website
IR.6	Success and success criteria	Success defined by Seamless interface with DTIC	Success defined by deposit rate and tracking hits
IR.7	Ensuring Success	Advocacy (Marketing) Institutional buy-in Keeping personnel up-to-date Keeping up with DTIC updates	Free, unfettered access Expanding to other languages

Table A-15: Responses to IR-Related Questions III

REFERENCES

- Air Force Institute of Technology, Public Affairs Office. (2006, May 25th). *AFIT Academic Library/About the Library*. Retrieved October 25th, 2008, from The Air Force Institute of Technology: <http://www.afit.edu/library/page.cfm?page=37>
- Alavi, M., & Leidner, D. E. (2001). Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. *MIS Quarterly*, 25 (1), 107-136.
- Allard, S., Mack, T. R., & Feltner-Reichert, M. (2005). The Librarian's Role in Institutional Repositories. *Reference Services Review*, 33 (3), 325-336.
- Andresen, D., Carver, L., Dolin, R., Fischer, C., Frew, J., Goodchild, M., et al. (1995). The WWW Prototype of the Alexandria Digital Library. *Proceedings of ISDL '95: International Symposium on Digital Libraries* (pp. 17-27). Santa Barbara: ISDL.
- Aquatic Commons. (2008). *About the Repository*. Retrieved February 15th, 2009, from Aquatic Commons: <http://aquacomm.fcla.edu/information.html>
- Ardichvili, A., Page, V., & Wentling, T. (2003). Motivation and Barriers to Participation in Virtual Knowledge-Sharing Communities of Practice. *Journal of Knowledge Management*, 7 (1), 64-77.
- Baase, S. (2008). *A Gift of Fire - Social, Legal, and Ethical Issues for Computing and the Internet* (3rd ed.). Upper Saddle River, New Jersey, USA: Pearson Prentice Hall.
- Bailey, C. W. (2007). *Institutional Repositories SPEC Kit*. Retrieved December 21st, 2008, from Digital Scholarship: <http://www.digital-scholarship.org/cwb/arlr.htm>
- Bailey, C. W. (2005). The Role of Reference Librarians in Institutional Repositories. *Reference Services Review*, 33 (3), 259-267.
- Barker, E. (1956). *From Alexander to Constantine*. (E. Barker, Trans.) Oxford, Oxford, UK: Oxford University Press.
- Bell, S., Foster, N. F., & Gibbons, S. (2005). Reference Librarians and the Success of Institutional Repositories. *Reference Services Review*, 33 (3), 283-290.
- Brown, J. S., & Duguid, P. (1998). Organizing Knowledge. *California Management Review*, 90-111.
- Buede, D. M. (2000). *The Engineering Design of Systems*. New York, New York, USA: John Wiley & Sons, Inc.

- Chambers, M. (1958). *Ancient Greece*. Washington, DC, USA: American Historical Association.
- Chambers, M., Hanawalt, B., Rabb, T. K., Woloch, I., & Grew, R. (1999). *The Western Experience* (7th ed.). Boston, Massachusetts, USA: McGraw-Hill College.
- Creswell, J. W. (2003). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (2nd ed.). Thousand Oaks, California, USA: Sage Publications, Inc.
- Crow, R. (2002). The Case for Institutional Repositories: A SPARC Position Paper. 37. Washington, D.C., USA: The Scholarly Publishing & Academic Resources Coalition.
- Davenport, T. H., & Prusak, L. (1998). *Working Knowledge*. Boston: Harvard Business School Press.
- Davis, P. M., & Connolly, M. J. (2007). Institutional Repositories: Evaluating the Reasons for Non-Use of Cornell University's Installation of DSpace. *D-Lib Magazine (Pre-Publication)* , 13 (3/4), 19.
- Dillman, D. A. (2007). *Mail and Internet Surveys: The Tailored Design Method* (2nd ed.). Hoboken, New Jersey, USA: John Wiley and Sons, Inc.
- DSpace Foundation. (2008, November 5th). *About DSpace*. Retrieved November 9th, 2008, from DSPACE: <http://www.dspace.org/>
- DSpace Foundation. (2007, June 13th). *End User FAQ*. Retrieved November 9th, 2008, from DSpace Wiki: <http://wiki.dspace.org/index.php/EndUserFaq>
- Einhard. (w. ca. aD 770-840; t. aD 1998). *Charlemagne's Courtier*. (P. E. Dutton, Ed., & P. E. Dutton, Trans.) Peterborough, Ontario, Canada: Broadview Press.
- EPrints. (2009). *EPrints - Digital Repository Software*. Retrieved March 1st, 2008, from Open Access and Institutional Repositories with EPrints: <http://www.eprints.org/software/>
- Gairdner, J. (1904). *A History of the English Church in the Sixteenth Century* (Vol. IV). London, United Kingdom: MacMillan and Co., Limited.
- Genoni, P. (2004). Content in Institutional Repositories: A Collection Management Issue. *Library Management* , 25 (6-7), 300-306.
- Gibbons, S. (2004, July-August). Establishing an Institutional Repository. (C. Santilli, Ed.) *Library Technology Reports* , p. 68.

Gray, D. (2003, November 1st). Wanted: Chief Ignorance Officer. *Harvard Business Review* , 22-24.

Green, P. (1993). *Alexander to Actium*. Berkeley, California, USA: University of California Press.

Hahn, J., Kannan, K., & Zhang, C. (2004). *Markets vs. Communities: An Empirical Analysis in the Knowledge Exchange Context*. Purdue University, Krannert School of Management. West Lafayette, Indiana: Working Paper; Purdue University.

Hammer, M., & Champy, J. (2003). *Reengineering the Corporation*. New York: Collins Business Essentials.

Hammer, M., Leonard, D., & Davenport, T. (2004). Why Don't We Know More About Knowledge. *MIT Sloan Management Review* , 14-18.

Hansen, M. T., & von Oetinger, B. (2001, March). Introducing T-Shaped Managers - Knowledge Management's Next Generation. *Harvard Business Review* , 106-116.

Heminger, A. R., & Robertson, S. B. (1998). Digital Rosetta Stone: A Conceptual Model for Maintaining Long-term Access to Digital Documents. *System Sciences, 1998., Proceedings of the Thirty-First Hawaii International Conference on System Sciences* , 2, 158-167 .

Henczel, S. (2000). The Information Audit as a First Step Towards Effective Knowledge Management: An Opportunity for the Special Librarian. *Worldwide Conference on Special Librarianship*. 34, pp. 210-226. Brighton: INSPEL.

Hoque, F. (2008). Competition from Emerging Markets. *Baseline* (088), 35-37.

Hylton, A. (2002, July). Measuring & Valuing Knowledge: The Role of the Knowledge Audit. *Measuring & Valuing Knowledge: The Role of the Knowledge Audit* , 9. Hylton Associates.

Ives, W., Torrey, B., & Gordon, C. (1998). Knowledge Management: An Emerging Discipline with a Long History. *Journal of Knowledge Management* , 1 (4), 269-274.

Johnson, R. K. (2002, October 28). Institutional Repositories: Partnering with Faculty to Enhance Scholarly Communication.

Kinder, H., & Hilgemann, W. (1974). *The Anchor Atlas of World History, Volume I* (Anchor Books ed., Vol. I). (E. A. Menze, Trans.) Garden City, New York, USA: Anchor Press/Doubleday.

Krippendorff, K. (2004). *Content Analysis: An Introduction to its Methodology*. Hershey, Pennsylvania, The United States of America: Sage Publications.

Langer, W. L. (1972). *An Encyclopedia of World History* (Fifth ed.). (W. L. Langer, Ed.) Boston, Massachusetts, USA: Houghton Mifflin Company.

Liebowitz, J., Rubenstein-Montano, B., McCaw, D., & Buchwalter, J. (2000). The Knowledge Audit. *Knowledge and Process Management* , 7 (1), 3-10.

Lyman, P., & Varian, H. R. (2000, October 18th). *How Much Information*. (University of California, Berkeley) Retrieved November 9th, 2008, from School of Information Management and Systems, UC Berkeley:
<http://www2.sims.berkeley.edu/research/projects/how-much-info/>

Lynch, C. A. (2003). Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age. *ARL* (226), 1-7.

Mackie, M. (2004). Filling Institutional Repositories: Practical Strategies from the DAEDALUS Project. *Ariadne* (39), 6.

MacLeod, R. (2005). *The Library of Alexandria*. (R. MacLeod, Ed.) London: I.B. Tauris Publishers.

Maxwell, J. A. (2004). *Qualitative Research Design: An Interactive Approach* (2nd ed.). Thousand Oaks, California, United States: Sage Publications, Inc.

Meyers, M. D., & Newman, M. (2006). The Qualitative Interview in IS Research: Examining the Craft. *Information and Organization* , 17, 2-26.

Neuman, W. L. (2006). *Social Research Methods: Qualitative and Quantitative Approaches* (6th ed.). Boston, Massachusetts, USA: Pearson Education, Inc.

Nonaka, I. (1994). A Dynamic Theory of Organizational Knowledge Creation. *Organization Science* , 5 (1), 14-37.

OCLC Online Computer Library Center, Inc. (2008). *About OCLC*. Retrieved December 12th, 2008, from OCLC Home: <http://www.oclc.org/us/en/about/default.htm>

OCLC Online Computer Library Center, Inc. (2009). *What is WorldCat*. Retrieved February 23rd, 2009, from WorldCat: <http://www.worldcat.org/whatis/default.jsp>

Osterloh, M., & Fey, B. S. (2000). Motivation, Knowledge Transfer, and Organizational Forms. *Organization Science* , 11 (5), 538-550.

Pattison, M. Q. (2002). *Qualitative Research & Evaluation Methods*. Thousand Oaks, California, United States of America: Sage Publications, Inc.

Plato. (w. 399 BC; t. aD 1997). *Defence of Socrates; Euthyphro; Crito*. (D. Gallop, Trans.) Oxford, Oxford, United Kingdom: Oxford University Press.

President and Fellows of Harvard College. (2007). *The Early History of Harvard University*. Retrieved January 25th, 2009, from The Harvard Guide: <http://www.hno.harvard.edu/guide/intro/index.html>

Prosser, D. (2003). Institutional Repositories and Open Access: The Future of Scholarly Communication. *Information Services and Use* , 23, 167-170.

Rosenwein, B. H. (2002). *A Short History of the Middle Ages*. Peterborough, Ontario, Canada: Broadview Press.

Schwab, D. P. (2005). *Research Methods fro Organizational Studies* (2nd ed.). (A. Duffy, Ed.) Mahwah, New Jersey, USA: Lawrence Erlbaum Associates, Publishers.

Shearer, K. (2003). Institutional Repositories: Towards the Identification of Critical Success Factors. *CAIS 2003: Bridging the Digital Divide: Equalizing Access to Information and Communication Technologies* (pp. 250-263). Halifax, Nova Scotia: Canadian Association for Information Science.

SHERPA, University of Nottingham. (2006). *About SHERPA*. Retrieved November 9th, 2008, from SHERPA: <http://www.sherpa.ac.uk/>

Smith, M. (2002). DSpace: An Institutional Repository from the MIT Libraries and Hewlett Packard Laboratories. (M. Agnosti, & C. Thanos, Eds.) *ECDL 2002, LNCS 2458* , 543-549.

The Berkeley Electronic Press™. (1999-2008). *Digital Commons*. Retrieved November 9th, 2008, from Berkeley Electronic Press: <http://www.bepress.com/ir/>

The Defense Technical Information Center. (2008, August 14th). *About DTIC*. Retrieved December 16th, 2008, from DTIC Online: <http://www.dtic.mil/dtic/aboutus/>

The Ohio Library and Information Network. (2008, October 3rd). *What is OhioLINK?* Retrieved February 15th, 2009, from OhioLINK: <http://www.ohiolink.edu/about/what-is-ol.html>

The Regents of The University of California. (2008). *A World of Digital Resources*. Retrieved November 9th, 2008, from Calisphere: <http://www.calisphere.universityofcalifornia.edu/>

The Regents of the University of California. (2008). *California Digital Library*. Retrieved November 9th, 2008, from California Digital Library: <http://www.cdlib.org/>

- The Regents of the University of California. (2009). *California Management Review*. Retrieved March 1st, 2009, from University of California Berkeley, Haas School of Business: http://cmr.berkeley.edu/subscriber_online_access.html
- The Regents of the University of California. (2008). *eScholarship Publishing*. Retrieved November 9th, 2008, from California Digital Publishing: <http://www.cdlib.org/programs/escholarship.html>
- The Regents of the University of California. (2006). *Melvyl - Basic Search*. Retrieved November 9th, 2008, from Melvyl: <http://melvyl.cdlib.org/>
- The Regents of the University of California. (2009). *OAI Harvesting Infrastructure*. Retrieved March 01, 2009, from Inside CDL: <http://www.cdlib.org/inside/projects/harvesting/>
- The Regents of the University of California. (2009). *Reshaping Scholarly Communication*. Retrieved March 1st, 2008, from The Office of Scholarly Communication: <http://osc.universityofcalifornia.edu/>
- The University of Cambridge. (2009, January). *800 Years of History*. Retrieved January 25th, 2009, from University of Cambridge: <http://www.800.cam.ac.uk/page/5/history.htm>
- The University of Oxford. (2008, September 16th). *A Brief History of the University*. Retrieved January 25th, 2009, from University of Oxford: http://www.ox.ac.uk/about_the_university/introducing_oxford/a_brief_history_of_the_university/index.html
- Utterback, J. M. (1994). *Mastering the Dynamics of Innovation*. Boston: Harvard Business School Press.
- Van de Sompel, H., Payette, S., Erickson, J., Logoze, C., & Warner, S. (2004). Rethinking Scholarly Communication - Building the System that Scholars Deserve. *D-Lib Magazine*, 10 (9).
- von Hippel, E. (1988). *The Sources of Innovation*. New York: Oxford University Press.
- Ware, M. (2004). Institutional Repositories and Scholarly Publishing. *Learned Publishing*, 17 (2), 115-124.
- Webster's New Millennium™ Dictionary of English, Preview Edition (v 0.9.7). (2008). *Google - Definitions from Dictionary.com*. Retrieved August 15th, 2008, from Dictionary.com: <http://dictionary.reference.com/search?r=2&q=Google>

Wikimedia Foundation, I. (2008, October 25th). *Great Libraries of the Ancient World*. Retrieved October 25th, 2008, from Wikipedia, the Free Encyclopedia:
http://en.wikipedia.org/wiki/Great_libraries_of_the_ancient_world

Yin, R. K. (2003). *Case Study Research: Design and Methods* (3rd ed., Vol. 5). Thousand Oaks, California, USA: Sage Publications.

VITA

Captain Alexander N. Constantine graduated from Monterey High School, Monterey California in June of 1999. After graduating High School, Capt Constantine attended the University of California, Berkeley, graduating in 2003 with a Bachelor's degree in Physics and an ROTC commission in the United States Air Force.

His first assignment was to the Propulsion Directorate of the Air Force Research Laboratory. Captain Constantine served as Chief of Projects and Flight Test Lead for the X-51A Scramjet Engine Demonstrator (SED) program, Propulsion Technology Branch, Aerospace Propulsion Division, in which capacity he led a multidisciplinary team of over 50 government and contractor personnel nationwide in organizing schedules, instrumentation, facilities, logistics, and the test bed for the X-51 flight test program. He was also responsible for developing and publishing the X-51 Security Classification Guide and maintaining the X-51 case file and electronic program information system.

While fulfilling his duties in the X-51 Program, he also competed into AFIT's Operational Technology and Innovation program, completing a graduate certificate while working on Project AngelFire – an air-based persistent ISR platform – developing user-interface technology and formulating/implementing test plans for AngelFire's highly successful pre-deployment exercise at the Marine Corps Air-Ground Combat Center, Twentynine Palms, California.

In August 2007, he entered the Graduate School of Engineering and Management, Air Force Institute of Technology to pursue MS degrees in Research and Development Management and Systems Engineering. Upon graduating, he will be assigned to Hill AFB, Utah, managing sustainment activities for the T-38 Talon.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 074-0188	
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of the collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p> <p>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p>					
1. REPORT DATE (DD-MM-YYYY)		2. REPORT TYPE		3. DATES COVERED (From – To)	
16-03-2008		Master's Thesis		August 2007-March 2009	
4. TITLE AND SUBTITLE Information Technology and the Evolution of the Library				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Constantine, Alexander N, Capt USAF				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way WPAFB OH 45433-7765				8. PERFORMING ORGANIZATION REPORT NUMBER AFIT/GRD/ENV/09M-01	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT The purpose of this research is to examine the current state of the practice of information technology use among academic and research libraries in order to gain insight into the level of maturity of various emerging technologies such as the Institutional Repository (IR) to be used to provide descriptive guidance and context to other institutions considering them for application. This research uses a case study methodology involving a single case design with multiple units of analysis, and draws evidence from individual interviews of research and education libraries chosen from a range of sizes, foci, and jurisdictions. This evidence is coupled with review of those institutions' public documentation and that of their partners to assess what academic libraries are doing today to meet the needs of the institutions they serve, what they have planned to provide these services, and what unplanned activities they feel are important to continue to make the information they steward available. Respondents cited several key information technology applications including institutional repositories, federated searching, and other integrated library services as tools being migrated towards to help them fulfill their missions. Interview and archival search results demonstrate a marked drive toward increased collaboration among institutions and regional repositories, a desire to provide rich content and seamless interfaces to library customers, and a trend towards hybrid systems to meet institutional archive and faculty communication needs.					
15. SUBJECT TERMS Knowledge Management, Library Science, Repository, Information Technology, Intellectual Property					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER (Include area code)
U	U	U	UU	134	Alan R. Heminger, PhD (937) 255-3636, ext 7405 e-mail: alan.heminger@afit.edu

